

有価証券報告書からの企業動向情報の抽出と可視化¹

長倉大輔研究会

武岡 弘明 ²

Dec. 30, 2018

概要

本論文では、投資家を対象に有価証券報告書から各企業の業務内容、直近の企業動向を計量的手法により簡略に可視化する方法を提案する。テキストマイニングにより有価証券報告書のテキストデータから名詞をすべて抽出し、出現回数などによりフィルタリング、ソートを行う。得られた頻出名詞を、その名詞を頻出名詞に含む全企業の時系列株価変化率を用いて「名詞ごとの評価変化率」を算定する。「名詞ごとの評価変化率」と「各企業の時系列株価変化率」の相関係数の大きさにより各企業の頻出名詞を並べ替えたものを各企業の出力結果とする。この出力結果を参照することにより、テキストリストにより簡潔に企業を表すことが可能となり有価証券報告書を探し業務に関する部分を探し読む手間を大幅に削減することができる。

¹ 本論文の内容に対する責任は全て筆者に帰する。

² 経済学部経済学科 4 年

1. はじめに

2008年以降、デフレ脱却に向けた政策金利の引き下げにより低金利が長期間続いている。超低金利の中、個人が資産形成する上で選択肢として株式投資や投資信託を選択する人が今後増えて行くことが予想される。米国では、税制優遇(IRA)などの政策対応により、家計でバランスのとれたポートフォリオが実現し、金融資産も大きく増加している。また、英国でもISA(Individual Savings Account: 個人貯蓄口座)が導入・恒久化され、投資額が増加している。日本銀行調査統計局の「資金循環の日米欧比較」によると2018年8月日本14.9%であるのに対し、アメリカは48%である。この日本の金融資産割合の少なさに対し政府も問題意識を抱えており、日本の家計における資産形成を促すために、NISA(Nippon ISA: 小額投資非課税制度)を通して株式投資を活性化させようとしている。2018年6月23日に金融庁が公表した「NISA口座の利用状況等について」によると、NISAの総口座数は650万3,951で制度導入時点の約475万口座から3カ月で約175万口座、率にすると約37%増加している。低金利が続く中、家計の資産形成手段として株式投資や投資信託の割合は、政策による促進と相まって一段と増えて行くはずである。

そういった現状であるにもかかわらず、投資家向けに企業情報として公開されている決算短信や有価証券報告書は分量が多い上に、会計的な単語が大部分を占めるため本業や注力している部分だけを抜粋し探して読んでいくのは個人投資家にとって大変である。これらのテキストデータには過去や将来の業績に対する理由や根拠、経営戦略に関する情報など数値データには含まれない重要な投資判断情報が含まれているため無視することはできない。

近年、膨大なデータベースや、自然言語解析、人工知能分野の手法を用いた文章解析などによる情報資源有効活用した投資の期待が高まりつつあり関連する論文が多数発表されている。Bollen, Mao and Zeng(2010)のようにツイッターのテキストデータから翌日のダウ・ジョーンズ工業株価平均の変動を予測する研究[1]や、日銀が発行する金融経済月報に対してテキストマイニングを行う和泉、後藤、松井(2011)の論文[3]などがあげられる。しかし個人投資を活性化させるという観点では、それらの論文は実用的でない。市況感全体を把握するような内容だからである。資産規模が大きい個人は、市場全体を踏まえたオートフォリオを組むことはできず、市場全体よりも各企業の事業内容、注力している業務に興味関心がある。

そこで本論文では、言語情報資源対し自然言語解析を用いることにより、シンプルに理解しやすい形まで成形し落とし込む手法について提案する。この成形をすること

により、企業情報を簡単に把握することができ、個人投資の参入を手助けすることにつながる。具体的には、有価証券報告書を形態素解析し、各企業の事業内容や直近力を入れている事業を端的に表す単語を計量的手法により抽出し可視化する手法を提示し、考察する。本稿では第2、3、4節で対象データ、抽出方法、抽出結果を示す。第5節は考察である。

2. 対象データ

本稿の手法では現時点で日経225構成銘柄として登録されている企業225個に対して大きく分けて二つの種類のデータを用意する。利用するデータの概要を表1に示す。テキストデータとして有価証券報告書を使い、数値データとしては株価を用いる。2.1,2.2節において有価証券報告書、株価それぞれの対象データについて詳しく述べる。

表1 利用データ

対象	内容
期間	2018年1月1日～2018年12月31日
企業	日経225構成企業
有価証券報告書	発表日ベースで対象期間に含まれている本決算の有価証券報告書
項目	「事業の内容」から「事業等のリスク」の前まで
株価	2018年1月4日～2018年12月28日までの株価終値245個

2.1 有価証券報告書

本論文では、金融商品取引法に基づき2018年1月1日から2018年12月31日までの期間に金融庁に提出された日経225構成銘柄の有価証券報告書225個を使用する。

2.1.1では他のテキストデータでなく有価証券報告書を抽出とした理由を述べ、2.1.2では有価証券報告書の中でさらに限定して利用する項目について説明する。

2.1.1 抽出対象とした理由

各企業を表す正式なテキストデータとして、決算短信、四半期報告書、有価証券報告書が公開されている中、本論文で有価証券報告書を選択した理由は佐藤、佐久間、小寺(2017)[2]で触れられているように他の二つのテキストデータに比べ企業を説明

する文章が多いからである。四半期報告書はそれまでに提出された最新の有価証券報告書に書いてある内容から変化がない場合、省略される部分が多い。特に投資家が注目する「事業の内容」などの項目は有価証券報告書を参照するように指示して省略されるケースが頻繁にある。また決算短信についても同様に事業内容は前提として変化があった近況についてのみ簡潔に述べられておりテキストデータ自体が豊富でない。以上の理由によりテキストデータが豊富であり、企業の業種、事業内容を省略せずに説明している有価証券報告書を利用した。

2.1.2 対象とする項目

有価証券報告書は分量が多く、報告上必要とされる会計的な内容が多く含まれているため、本稿では使用する項目を限定する。有価証券報告書は記載事項の大枠項目は金融商品取引法で定められており、大項目として1.企業情報、2.提出会社の保証会社等の情報、3.監査報告書がある。企業の事業内容が文章として書かれているのは1の企業情報であり、さらにその中で直接的にこれまでの事業内容や業績、今後注力していく内容が書かれている部分は、企業情報の中項目である「企業の概況」内の「事業の内容」から、同じく中項目である「事業の状況」内の「事業等へのリスク」の前までである。よって本論文では「事業の内容」から「事業等へのリスク」を利用した。

2.2 株価

本論文では有価証券報告書のテキストデータと併せて、数値データとして各日経225構成銘柄に対し2018年1月4日から2018年12月28日までの株価終値245個を使用する。2.1で述べたように有価証券最新の有価証券報告書の内容にちょうど対応する期間は各企業によって異なるが、本論文では日経225構成銘柄の企業を横断して単語ごとの変化率を生成するため、株価のデータ期間を統一する必要がある。そのため2018年12月31日時点までに提出された最新の有価証券報告書に対応する株価データを、2018年1月4日から2018年12月28日までの終値とした。

3. 抽出方法

抽出方法の大きな流れとしては、まず銘柄225個一つ一つの有価証券報告書の該当項目に対し形態素解析を行い、名詞をすべて抽出する。数字や記号、直接業務内容を

表すと考えられない名詞を排除し、各企業の頻出名詞のリストを作る。(3.1) この名詞リストの各名詞に対し、対数階差により算出した株価変化率をもとに名詞ごとに評価変化率を求める。(3.2) この各名詞の評価変化率とその名詞を含む企業の株価変化率の相関係数をさらに求め、相関係数が大きい順に各企業の頻出名詞リストを並べ替える。(3.3) これにより、各企業に対し株価変化率と相関が強い順に頻出名詞リストを並べ直すことができ、企業が注力している事業の概要をつかみやすいように成形することができる。

この抽出方法の過程において、以下に示す調整可能なパラメータが存在する。これら3つの値 (α 、 β 、 γ) を変化させることにより個人投資家は各々出力結果を調整することができる。変数 α は3.1.2の手順2、変数 β は3.1.3の手順3、変数 γ は3.3において詳しく説明される。本論文では一例として、 $\alpha=4$ 、 $\beta=200$ 、 $\gamma=5$ の値を選択する。

[変数 α ：頻出名詞として扱う上での最低出現数] (3.1.2の手順2を参照)

[変数 β ：余計な名詞として扱う上での最低出現数] (3.1.3の手順3を参照)

[変数 γ ：ある名詞の相関係数を0とする出現企業数] (3.3を参照)

3.1 各企業における名詞リストの作成方法

日経225構成銘柄の銘柄コード225個の一つ一つに対し以下の操作をする。

3.1.1 各企業名詞リストの用意

手順1：

有価証券報告書のPDFをjavaのライブラリであるpdfboxを用いてテキスト化する。

この際表が含まれていた場合は表に含まれるテキスト情報もテキスト化されてしまうが後に処理によりその影響は除外される。

手順2：

テキスト化された有価証券報告書全体の中から、「事業内容」と「事業等へのリスク」という言葉の位置を探し、その間のテキストだけ抽出する。

手順3：

該当箇所だけになったテキストを句点で区切り、各文をjavaの自然言語処理のライブラリであるkuromojiにより形態素解析する。kuromojiにより名詞と判断された名詞を集め、「名詞リストA」とする。

3.1.2 各企業名詞リストの篩分け1

手順1：

まず数字や記号だけで名詞と判断されてしまった単語は、正規表現を利用することにより名詞リスト A から省く。これを「名詞リスト B」とする。

省いた数字や記号は以下の正規表現に一致するものである

[0-9 0-9 -- !"# \$%&' () =~|' {+*} <>? _- ^ ¥@ 「 ; : 」 , .
①②③④⑤⑥⑦⑧⑨ 【 】]+]

手順2：

名詞リスト B の名詞で重複しているものをカウントし、カウント数が多いものから順に降順で並べる。カウント数 α 回以上重複しなかったものを省く。これを「名詞リスト C」とする。

重複する名詞をすべて表示すると量が多くなり視認性が下がるため、「一つの有価証券報告書内で α 回以上重複すること」をその企業の現状をよく表す言葉の条件とした。これにより大幅に名詞リストの数を減らし投資家の視認性をあげることにつながる。

3.1.3 業務を直接表さない余計名詞リストの作成

上述したように名詞の篩分けを行い名詞リスト C の段階に至っても、上位の名詞には「事業」や「株式」、「会社」、「円」などの有価証券報告書を記述する上で説明上避けられない名詞や、会計上で必要な名詞が紛れてしまう。このような業務内容を直接理解できない名詞を省き、事業内容の概要を単語から推察できるようにすることが本論文の目的である。そのため、頻出する余計な名詞を別にリスト化（「余計名詞リスト」とする）し、各企業の名詞リスト C から排除する手順を踏む。日経 225 構成銘柄の有価証券報告書全体に出てくる名詞で頻出のものに注目し余計名詞リストを作成する。

手順1：

まず日経 225 に含まれる各企業に対し、3.1.1,3.1.2 の手順 1 に従い名詞リスト B までを作成する。

手順 2 :

手順 1 で作成された 225 個の名詞リスト B をすべて結合する。これにより数字や記号を含まない有価証券報告書に出てくる名詞すべてが入ったリスト（「名詞リスト X」とする）ができる。

この名詞リスト X における頻出回数上位の単語は有価証券報告書全体を通してよく出てくる単語となり、今回省きたい余計な名詞が含まれているととらえられる。

手順 3 :

名詞リスト X に出てくる名詞の出現数をカウントし、 β 回以上出現したものを抽出し「名詞リスト Y」とする。

この時点では名詞リスト Y の段階では各企業の業務を表す可能性があるものや影響を受ける内容を端的に表せそうな単語も含まれているため、このまま全体を通して出現回数が β 回以上ものをすべて余計単語リストとみなすのは問題である。そのため追加的に手順 4 を行う。

手順 4 :

「企業の業務を表す可能性がある名詞」や「企業が影響を受ける内容（地域名、地名）を端的に表しそうな名詞」を名詞リスト Y から分析者が直接視認して判断し、名詞リスト Y から排除する。これを最終的な余計名詞リストとする。

以下の表 2 が名詞リスト Y から除外した企業の業務を表しうると判別し選出した名詞である。最終的な余計名詞リストは論文末に表 3 として示した。

表 2 業務に直接関係しうる名詞 ($\alpha=4$, $\beta=200$)

製造	通信	資源	輸送	人材	セメント
東京	不動産	材料	設計	アメリカ	電力
サービス	米国	食品	石油	金属	工業
開発	金融	物流	住宅	派遣	シンガポール
情報	建設	研究	医療	発電	インフラ
機械	電子	工事	船	原料	台湾
中国	電気	医薬品	ブランド	タイ	原材料
機器	保険	大阪	運用	半導体	生命
システム	エネルギー	銀行	アジア	土地	繊維
部品	港	化学	上海	ユーロ	鉄道
自動車	千代田	ガス	欧州	プラント	データ

3.1.4 各企業名詞リストの篩分け 2

「3.1.2 各企業名詞リストの篩分け 1」で生成された名詞リスト C から「3.1.3 業務を

直接表さない余計名詞リストの作成」で最終的に生成された余計名詞リストを除く。余計名詞を排除した最終名詞リストを「頻出名詞リスト」とする。

以上 3.1.1 から 3.1.4 により各企業に対し、業務内容や業務に影響を及ぼす名詞だけを集めた最終的な名詞リストが作成された。この最終名詞リストを「頻出名詞リスト」とする。

3.2 名詞ごとの評価変動率算定

2018 年 1 月 4 日から 2018 年 12 月 28 日までの企業 A の株価終値 245 個に対し、株価変化率 $S_A = (s_2, s_3, \dots, s_{245})$ の各成分は(1)式のように各時系列データの対数階差を取ったものとする。

$$s_t = \Delta \log y_t = \log(y_t) - \log(y_{t-1}) = \log\left(\frac{y_t}{y_{t-1}}\right) = \log\left(1 + \frac{y_t - y_{t-1}}{y_{t-1}}\right) \approx \frac{y_t - y_{t-1}}{y_{t-1}} \quad (1)$$

ここで y_t は第 t 日の株価終値である。

各企業について名詞リストを並べ替えるため、名詞ごとに評価変化率を算出する。本論文で提案する任意の名詞（名詞 a）について評価変化率を算出する手順は以下のようである。

手順 1 :

3.1 で最終的に作成された各企業の頻出名詞リストに名詞 a を含む企業を探し、該当企業の集合を集合 C とする

手順 2 :

集合 C に含まれるすべての企業の株価変化率に対し、同時点の成分をすべて足し平均を取る。これにより得られた 244 次元の変化率平均値を名詞 a の評価変化率 W_a とする。

$$W_a = \frac{1}{n} \sum S_c, c \in C, n = |C| \quad (2)$$

3.3 相関係数による名詞リストの並べ替え

ある企業 A の頻出名詞リストに含まれる名詞を、「名詞の評価変化率」と「企業 A の株価変化率」の相関係数が大きい順に並べ替える。これにより企業 A の株価変化率と相関が強い名詞の順に頻出名詞リストを並べ直すことができ、企業 A が注力している事業の概要をつかみやすいように成形できる。

「名詞 a の評価変化率 W_a 」と「名詞 a を含む企業 A の株価変化率 S_A 」の相関係数は以下の式(3)で求める。

$$\text{cor}(S_A, W_a) = \frac{\sum_{t=2}^{245} (s_t - \bar{s})(w_t - \bar{w})}{\sqrt{\sum_{t=2}^{245} (s_t - \bar{s})^2} \sqrt{\sum_{t=2}^{245} (w_t - \bar{w})^2}} \quad (3)$$

$$S_A = (s_2, s_3, \dots, s_{245}), W_a = (w_2, w_3, \dots, w_{245}),$$

$$\bar{s} = \frac{1}{244} \sum_{t=2}^{245} s_t, \bar{w} = \frac{1}{244} \sum_{t=2}^{245} w_t$$

この頻出名詞リストを相関係数により並べ替える処理を日経 225 構成銘柄すべてに対して行う。ただし、企業特有の名詞で他の企業に含まれていない名詞が頻出名詞リストに含まれていた場合、評価変化率と該当企業の株価変化率の相関係数は極めて高くなってしまう。こういった名詞がすべて上位になると一般的に知られていない名詞が上位を占める可能性があり解釈性が低下する。解釈性の低下を防ぐため、本論文では 3.2 の手順 1 でリストアップされた企業数が γ 個以下の場合、株価変化率と名詞変化率の相関係数を 0 とした。

4. 抽出結果

$\alpha=4$, $\beta=200$, $\gamma=5$ のもと、有価証券報告書に出現する頻出名詞リストを並び替えた結果上位 10 個を示したものが論文末の表 4.1~4.5 である。数が多いため、銘柄コードにより 5 つの表に分けて表示した。表 4.1 は銘柄コード 1332~4188 に対応しており、図表 4.2 は銘柄コード 4208~6103、表 4.3 は銘柄コード 6113~7270、表 4.4 は銘柄コード 7272~9009、表 4.5 は銘柄コード 9020~9984 に対応している。

5. 考察

抽出結果が各企業の事業概要を表せているか確認し、抽出結果に対して考察する。2018 年 12 月 31 日時点で国内時価総額上位 10 位に入る日経 225 構成銘柄のうち、業種が異なる 5 つの企業（トヨタ、NTT ドコモ、三菱 UFJ 銀行、ソニー、ファーストリテイリング）の抽出結果を以下の図表 6 として示す。図表 7 は同じ企業 5 つに対し抽出過程で出力される頻出名詞リストである。

表 5.1 時価総額上位企業の抽出結果

番号	企業名	頻出名詞リストのうち相関係数上位10個の名詞
7203	トヨタ自動車(株)	愛知:ファイナンス:名古屋:別:米国:自動車:中国:タイ:製造:設立
9437	(株)NTTドコモ	役務:とき:移動:合併:規定:範囲:パートナー:違反:適正:使用
8306	(株)三菱UFJフィナンシャル・グループ	金利:信託:リテール:各社:期待:高度:個人:提携:ニューヨーク:三菱
6758	ソニー(株)	マーケティング:兆:半導体:量:損害:中国:製造:エレクトロニクス:事務所:サービス
9983	(株)ファーストリテイリング	出店:プロジェクト:上海:米国:中国:情報:開発:ブランド:ユニ:クロ

表 5.2 時価総額上位企業の頻出名詞リスト

番号	企業名	頻出名詞リスト上位10個
7203	トヨタ自動車(株)	自動車:トヨタ:トヨタ自動車:開発:部品:モーター:金融:愛知:研究:製造
9437	(株)NTTドコモ	通信:電気:とき:接続:スマート:ライフ:大臣:ドコモ:省令:指定
8306	(株)三菱UFJフィナンシャル・グループ	銀行:三菱:UFJ:東京:本部:金融:信託:フィナンシャル:運用:一体
6758	ソニー(株)	ソニー:事務所:音楽:サービス:製造:東京:IP:半導体:映画:転借
9983	(株)ファーストリテイリング	ユニ:クロ:ジーユー:ブランド:ファーストリテイリング:迅:銷:中国:商:貿

まず抽出結果から各企業の業種が分かるかについて確認する。日本経済新聞による業種分類によると、トヨタ自動車は「自動車」、NTTドコモは通信サービス、三菱UFJフィナンシャルグループは「銀行」、ソニーは「家庭用電気機械」、ファーストリテイリングは「衣料品・服飾品」として分類されている。図表 6.1 を参照することにより「自動車」が含まれるトヨタ自動車と「エレクトロニクス」、「半導体」が含まれるソニーについては業種が分かる一方、NTTドコモ、三菱UFJ、ファーストリテイリングについては上位 10 個からは直接的に事業内容を読み取ることができない。しかし、抽出過程で出力される頻出名詞リストを見るとファーストリテイリング以外業種を直接表す言葉が入っているため、結果の上位には含まれていないが結果には含まれていることがわかる。この状況は各名詞を検索した企業数の少なさ起因すると考えられ、東証上場銘柄全体などより多くの銘柄に対して本論文と同じ手法を使えば業種分類と直接結び付く名詞が抽出結果の上位に来ると考えられる。

本論文は対象期間を 2018 年 1 月 4 日から 2018 年 12 月 28 日までとしているため、その期間内の変動に関係する名詞を抽出結果がとらえられているのか確認する。確認方法としては東洋経済新報社が発行している四季報（2018 年 4 集）の「材料記事」を利用する。トヨタ自動車は四季報の材料記事によると、「北米で奨励金が増えるがアジアや欧州堅調。」とあり、図表 6.1 の抽出結果にはタイ、中国が上位に表示されておりこの期間をよく表せる名詞を上位に並び替えることができていると分かる。またNTTドコモは材料記事[4]に「非通信は金融や決済中心に好進し事業拡大への積極投資こなす。」とあり、抽出結果の「合併」がこれを表していると考えられる。しかしNTTドコモは抽出結果の上位に直接業務と関係ない名詞が多く含まれている。これは手順

の変数である α を減らし余計名詞リストを充実させることにより改善できる。三菱UFJ フィナンシャルグループに関しては、四季報材料記事[4]に「外貨貸し出し続伸だが、円貸し出しは利ザヤ続落。役務は伸び悩む。」とあり、抽出結果上位の「金利」の影響を大きく受けていることがわかる。ソニーも材料記事[4]の「半導体は先行負担重く子会社譲渡益ないがスマホ用拡大で減益幅縮小。」に抽出結果の「半導体」が対応している。「ファーストリテイリング」も材料記事[4]によると「中国、東南アジアが牽引。」「欧米などでブランド認知拡大へ」という記事が抽出結果の「上海」、「中国」と「ブランド」がそれぞれあらわしている。

ここまで考察で紹介した企業は一部であり、他企業に対しても同様に各企業のテーマや動向の一部に関係する名詞を一覧で可視化することができているとは言いきれない。しかし、この抽出結果は個人投資家が銘柄を総覧し、短時間で企業の概要や直近で注力している事業の概要をつかめる点で大きく意義がある。本手法の欠点は各名詞が企業に対しポジティブな影響を与えたのか、ネガティブな影響を与えているのか判断することができないことである。この点を解消するためには本論文の手法に加えて文全体を自然言語解析しポジティブさ、ネガティブさを判定してくれる Opinion Finder[5]の利用などが考えられる。

表3 余計単語リスト ($\alpha=4, \beta=200$)

事業	ため	推進	供給	皆様	債務	組織	連携
当社	所有	現在	確保	ホールディングス	援助	結果	設置
等	数	議決	委員	世界	末	ビジョン	省
販売	会	法	米ドル	期間	提案	リスク	法人
(株)	—	日本	目標	構成	前期	経常	追求
会社	方針	有限	委託	間	社員	生活	費
円	成長	公司	体制	運営	国	業績	開催
グループ	セグメント	基本	持続	可能	財務諸表	4月	ガバナンス
万	割合	実現	中央	取得	維持	実績	製
経営	もの	課題	中	他	需要	整備	金額
連結	強化	以下	合計	変化	積極	信頼	歳
株式会社	付	(株)、	資産	さま	施策	—	項
百	営業	市場	購入	各種	プラン	賃借	実行
的	有	内	比	対処	向け	保有	既存
役員	売上	国内	必要	受託	店	推移	店舗
製品	権	労働	展開	買収	創造	消費	特別
こと	億	注	住所	全社	率	減少	賞与
兼任	向上	次	新た	臨時	創出	工場	材
年	平成	年間	財務	お客	該当	新規	持株
区	主要	員数	以上	増	中長期	発動	対象
提出	人	同社	基準	予約	契約	評価	最大
従業	買	3月	上記	社外	年数	もと	具体
証券	平均	活動	先	安全	条	対抗	センター
子会社	社会	規模	行為	保証	総会	策定	ビル
報告	取締役	賃貸	継続	就業	検討	物	支援
市	投資	産業	額	貸付	間接	づくり	ソリューション
書	金	とおり	出向	支配	長期	装置	統括
有価	提供	一部	部門	決定	内部	前年	代理
千	業	上	貢献	キャッシュ	影響	意思	最適
社	収益	判断	月	新株	保守	導入	系統
者	県	対応	よう	将来	施設	はじめ	ネットワーク
関係	設備	事項	増加	理念	損益	相互	モデル
利益	管理	株式	当期	これら	全体	マネジメント	賃金
都	資金	中期	州	ほか	指標	回復	確立
企業	名称	主	活用	重要	国際	加速	調達
その他	取引	品	お客様	出資)—	受注	ステーキ
内容	技術	分野	構築	品質	欄	利用	(株) (
員	高	外	建物	株	大量	適切	執行
関連	用	持分	給与	賃貸借	期	監査	ホルダー
書類	記載	ビジネス	特定	顧客	勤続	発展	更
業務	拡大	場合	取り組み	競争	構造	決議	総務
株主	資本	地域	これ	目的	エンジニアリング	雇用	事務
名	会計	人員	うち	独立	措置	二	
年度	商品	適用	領域	今後	多様	図	
状況	力	加工	経済	改善	策	概要	
日	戦略	基盤	中心	変更	総合	革新	
環境	計画	共同	効率	共通	後	所	
性	海外	当該	機	改革	ニーズ	コスト	
化	生産	機能	実施	フロー	達成	年齢	
価値	組合	グローバル	安定	区分	型	開始	

表 4.1 頻出名詞リストのうち相関係数上位 10 個の名詞(1332~4188)

番号	企業名	頻出名詞リストのうち相関係数上位10個の名詞
1332	日本水産(株)	健康:食品:物流:東京:研究:情報:資源:原料:水産:短期
1333	マルハニチロ(株)	冷凍:食:食品:ブランド:物流:商事:東京:北米:製造:マルハニチロ
1605	国際石油開発帝石(株)	石:種類:天然:分の:共和:石油:西:能力:5月:炭素
1721	コムシスホールディングス(株)	締結:インフラ:東京:建設:情報:港:システム:品川:通信:日本コムシス
1801	大成建設(株)	発注:建築:負債:土木:不動産:新宿:建設:東京:開発:工事
1802	(株)大林組	徹底:発注:施工:カナダ:水準:建築:土木:神奈川:方:制度
1803	清水建設(株)	発注:徹底:施工:財:着実:外部:安心:コンプライアンス:重点:建設
1808	(株)長谷工コーポレーション	企画:戸:住宅:圏:アメリカ合衆国:マンション:程度:分譲:不動産:都市
1812	鹿島	アジア:北米:工事:場:設計:発注:港:インターナショナル:投融資:土木
1925	大和ハウス工業(株)	戸:マンション:北:コア:住宅:時間:不動産:建築:リース:工業
1928	積水ハウス(株)	戸:北:マンション:分譲:住宅:建築:都市:施工:拡販:順調
1963	日揮(株)	横浜:西:神奈川:インフラ:化成:石油:プラント:建設:プロジェクト:高度
2002	(株)日清製粉グループ本社	健康:用地:安心:冷凍:事務所:食:無償:割当て:株券:勧告
2269	明治ホールディングス(株)	スポーツ:食:医薬:健康:コア:安心:薬:医薬品:千葉:食品
2282	日本ハム(株)	未来:視点:様々:フード:要件:転籍:尊重:意見:食:輸入
2432	(株)ディー・エヌ・エー	インターネット:スポーツ:充実:コンプライアンス:渋谷:東京:サービス:広告:運用:マーケティング
2501	サッポロホールディングス(株)	飲料:水準:出店:テーマ:飲食:機会:発売:減:償却
2502	アサヒグループホールディングス(株)	当年度:飲料:表示:類:重点:発売:数量:減:償却:拡充
2503	キリンホールディングス(株)	飲料:当年度:類:優先:伸長:順調:医薬:同等:範囲:数量
2531	宝ホールディングス(株)	バランス:イ:相当:交付:一定:卸売:買付け:諮問:5月:最大限
2768	双日(株)	程度:金属:航空:インフラ:電力:現地:化学:プロジェクト:リテール:都市
2801	キッコーマン(株)	飲料:健康:ジャパン:選任:旨:株券:食:いずれ:卸売:千葉
2802	味の素(株)	冷凍:トップ:健康:食:減:時間:商号:料:ベトナム:サポート
2871	(株)ニチレイ	指導:類:事務所:食品:満足:貿易:不動産:十分:自己:研究
2914	J T	テーマ:伸長:トップ:冷凍:食:当年度:満足:水準:バランス:フード
3086	J. フロント リテイリング(株)	出店:皆さま,:費用:表示:項目:作成:同等:9月:固定
3099	(株)三越伊勢丹ホールディングス	小売:方:最大限:百貨店:新宿:統合:不動産:専門:中華人民共和国:東京
3101	東洋紡(株)	マテリアル:繊維:フィルム:商事:それ:設定:北:重点:樹脂:防衛
3103	ユニチカ(株)	拡販:繊維:注力:フィルム:貿易:ガラス:京都:樹脂:中国:インドネシア
3105	日清紡ホールディングス(株)	精密:規制:燃料:素材:銅:繊維:ケミカル:化学:中国:エレクトロニクス
3289	東急不動産ホールディングス(株)	マンション:ホテル:都市,:渋谷:住宅:貸借:次世代:東京:開発
3382	(株)セブン&アイ・ホールディングス	出店:小売:減:沖繩:ジャパン:渋谷,:税:千葉:役務
3401	帝人(株)	輸出入:岡山:マテリアル:フィルム:繊維:樹脂:貸与:広島:化成:複合
3402	東レ(株)	繊維:南:フィルム:複合:炭素:化成:分の:合成:ケミカル:樹脂
3405	(株)クラレ	繊維:フィルム:樹脂:炭素:岡山,:南:公正:北:徹底
3407	旭化成(株)	マテリアル:姿:表示:繊維:コンプライアンス:行動:挑戦:ケミカル:財:安心
3436	(株)SUMCO	半導体:至:適正:原材料:同期:要因:台湾:価格:1月:売却
3861	王子ホールディングス(株)	家庭:岡山:包装:交付:期待:設立:法令:ハンド:提携:家
3863	日本製紙(株)	家庭:皆さま:以外:転換:家:ルール:土木:条件:無償:割当て
4004	昭和電工(株)	業界:炭素:景気:アルミ:上昇:合成:それぞれ:増益:半導体:樹脂
4005	住友化学(株)	供与:農業:フィルム:液晶:原材料:化学:住友:ニューヨーク:ライフ:カンパニー
4021	日産化学(株)	農業:医薬品:卸売:剤:薬:化学:開発:東京:材料:日産
4042	東ソー(株)	樹脂:上昇:ガラス:工業:原材料:化学:能力:石油:製造:原料
4043	(株)トクヤマ	特殊:フィルム:樹脂:広島:制度:原材料:化成:中華人民共和国:工業:テック
4061	デンカ(株)	供与:包装:資材:蘇州:特殊:工業:樹脂:完成:検査:商事
4063	信越化学工業(株)	合成:樹脂:半導体:台湾:能力:化成:商事:化学:米:国:製造
4151	協和発酵キリン(株)	順調:同等:作成:増減:譲渡,:治療:項目:医薬:償却
4183	三井化学(株)	工業:素材:人民元:化学:上海:台湾:中国:三井:シンガポール:製造
4188	(株)三菱ケミカルホールディングス	樹脂:三菱:素材:化学:コア:製造:ケミカル:ネット:大阪:医療

表 4.2 頻出名詞リストのうち相関係数上位 10 個の名詞(4208~6103)

番号	企業名	頻出名詞リストのうち相関係数上位10個の名詞
4208	宇部興産(株)	三菱:合成:樹脂:資材:不適切:体:工業:カンパニー:防止:化学
4272	日本化薬(株)	液晶:供与:ディスプレイ:アメリカ合衆国:樹脂:メキシコ:線:中華人民共和国:請負:責任
4324	(株)電通	アメリカ合衆国:ロンドン:マーケティング:時間:費用:帰属:減:償却:親会社:データ
4452	花王(株)	費用:伸長:増減:順調:表示:一定:兆:発売:減:一般
4502	武田薬品工業(株)	製薬:連合:疾患:医薬品:米国:研究:製造:開発:単位:武田
4503	アステラス製薬(株)	製薬:医薬品:医療:研究:コア:行動:ヨーロッパ:米国:台湾:ユーロ
4506	大日本住友製薬(株)	疾患:症:がん:細胞:メディカル:医薬:試験:治療:再生:剤
4507	塩野義製薬(株)	創:患者:医薬品:薬:剤:医療:治療:健康:症:拡充
4519	中外製薬(株)	創:治療:剤:薬:固定:患者:一時:税:製薬:医薬
4523	エーザイ(株)	患者:臨床:発行:法令:疾患:号:製薬:違反:試験:メキシコ
4543	テルモ(株)	メディカル:患者:治療:ベトナム:医療:ヨーロッパ:テック:コーポレート:現場:上海
4568	第一三共(株)	臨床:細胞:患者:疾患:症:制度:承認:試験:ジャパン:がん
4578	大塚ホールディングス(株)	治療:細胞:バランス:症:臨床:がん:伸長:飲料:発売:薬
4631	D I C(株)	埼玉:液晶:景気:合成:繊維:樹脂:上昇:増益:製造:中国
4689	ヤフー(株)	解決:インターネット:広告:親会社:メディア:情報:サービス:東京:港:コマース
4704	トレンドマイクロ(株)	スマート:現金:インターネット:同期:台湾:支出:研究:それ:様々:増収
4751	(株)サイバーエージェント	スマート:9月:インターネット:広告:育成:渋谷:メディア:東京:サイバーエージェント:ゲーム
4755	楽天(株)	至:項目:満足:会員:インターネット:1月:ペース:売却:獲得:カード
4901	富士フイルムホールディングス(株)	商号:ドイツ:愛知:バイオ:ディスプレイ:メディカル:プリンター:医療:神奈川:再生
4902	コニカミノルタ(株)	プリンター:貸与:複合:ユーロ:印刷:料:中国:医療:オフィス:製造
4911	(株)資生堂	米州:一方:累計:費用:譲渡:固定:地球:マーケティング:増減:ジャパン
5002	昭和シェル石油(株)	石:家庭:系:台:石油:要因:能力:9月:増益:融資
5020	J X T Gホールディングス(株)	石油:非鉄:天然:リサイクル:銅:鉱山:一(株):統合:横浜:プロジェクト
5101	横浜ゴム(株)	拡販:輪:強み:元:中華人民共和国:独自:工業:ロシア:自動車:タイヤ
5108	(株)ブリヂストン	輪:同等:バス:拡充:現金:重点:ユーロ:工業:米国:減
5201	A G C(株)	ガラス:景気:通り:ディスプレイ:私:上昇:化学:たち:融資:建築
5202	日本板硝子(株)	ガラス:通り:帰属:建築:米国:自動車:製造:親会社:削減:変革
5214	日本電気硝子(株)	ディスプレイ:ガラス:モノ:至:1月:プロセス:堅調:現金:建築:米国
5232	住友大阪セメント(株)	電池:セメント:建材:光:府:電子:工事:千代田:材料:物流
5233	太平洋セメント(株)	埼玉:着実:姿:江東:広島:用地:工業:神奈川:建築:事務所
5301	東海カーボン(株)	業界:同期:堅調:収入:価格:支出:工業:土地:中国:米国
5332	T O T O(株)	地球:ハンド:テーマ:マーケティング:諮問:。:米州:選任:オセアニア:財
5333	日本ガイシ(株)	供与:蘇州:制度:元:中華人民共和国:愛知:ユーロ:中国:米国:原材料
5401	新日鐵住金(株)	鉄鋼:熱:鋼板:素材:特殊:金属:亜鉛:資源:線:ブラジル
5406	(株)神戸製鋼所	熱:特殊:鉄鋼:鋼板:神戸:アルミ:航空機:教育:私:現場
5411	ジェイ エフ イー ホールディングス(株)	鉄鋼:鋼板:特殊:輸出入:非鉄:金属:亜鉛:調査:融資:火力
5541	大平洋金属(株)	現場:鉱山:公正:手続:コンプライアンス:対策:資源:価格:原料:製造
5631	(株)日本製鋼所	鋼板:修理:。:広島:担当:満足:早期:検査:樹脂:商事
5703	日本軽金属ホールディングス(株)	強み:化成:熱:別:。:アルミ:素材:中華人民共和国:体:静岡
5706	三井金属	亜鉛:股:銅:三井:金属:上海:計:自動車:製造:台湾
5707	東邦亜鉛(株)	亜鉛:鉱山:非鉄:リサイクル:金属:資源:原料:中国:製造:運輸
5711	三菱マテリアル(株)	リサイクル:教育:アルミ:銅:5月:金属:三菱:諮問:手続:相当
5713	住友金属鉱山(株)	鉱山:非鉄:銅:アメリカ合衆国:金属:マテリアル:テック:フィリピン:計:カナダ
5714	D O W Aホールディングス(株)	拡販:リサイクル:岡山:亜鉛:銅:名古屋:金属:南:愛知:締結
5801	古河電気工業(株)	電装:コーポレート:ガバナンス:神奈川:エレクトロニクス:原材料:インフラ:台湾:人材:自動車:製造
5802	住友電気工業(株)	住友:蘇州:人民元:素材:拡販:注力:電装:車両:線:インク
5803	(株)フジクラ	電装:原材料:商事:エレクトロニクス:上海:カンパニー:自動車:中国:情報:不動産
5901	東洋製罐グループホールディングス(株)	包装:客観:株券:手続:最終:政策:役:確認:原則:料
6098	(株)リクルートホールディングス	旅行:発行:プラットフォーム:飲食:広告:サポート:派遣:ユーロ:ドル:専門
6103	オークマ(株)	上海:機械:中国:製造:高度:部品:サービス:工作:アフターサービス:オークマ

表 4.3 頻出名詞リストのうち相関係数上位 10 個の名詞(6113~7270)

番号	企業名	頻出名詞リストのうち相関係数上位10個の名詞
6113	(株)アマダホールディングス	カナダ:神奈川:ブラジル:メキシコ:修理:ユーロ:中華人民共和国:モノ:上海:ロシア
6178	日本郵政(株)	締結:信託:生命:金銭:発揮:規制:自己:各社:公共:保険
6301	コマツ	種類:別:ブラジル:鉱山:ファイナンス:リテール:現場:ユーロ:機械:業界
6302	住友重機械工業(株)	精密:住友:機械:注力:中国:船舶:米国:人民元:製造:人材
6305	日立建機(株)	機械:中国:オーストラリア:行動:製造:アメリカ:部品:開発:建設:サービス
6326	(株)クボタ	農業:別:元:合成:エンジン:前年度:当年度:カナダ:水:拡充
6361	存原	貸与:股:別:前年度:精密:テック:水:プロセス:公共:石油
6366	千代田化工建設(株)	系:フィリピン:横浜:㈱(担当:インターナショナル:インドネシア:リミテッド:派遣:コーポレーション
6367	ダイキン工業(株)	蘇州:アメリカ合衆国:共和:冷凍:中華人民共和国:中国:インク:ヨーロッパ:化学:徹底
6471	日本精工(株)	精密:公正:ユーロ:担当:機械:それ:中国:設定:韓国:諮問
6472	N T N(株)	精密:担当:公正:中華人民共和国:上海:資源:中国:コンプライアンス:運転:自動車
6473	(株)ジェイテクト	愛知:ユーロ:政策:機械:中国:製造:アメリカ:部品:器具:工作
6479	ミネベアミツミ(株)	半導体:発揮:さまざま:精密:ドイツ:中国:製造:不適切:米国:機械
6501	(株)日立製作所	制御:イノベーション:デジタル:単位:空調:インド:超過:電子:通信:調整
6503	三菱電機(株)	種類:別:貸与:韓国:三菱:企画:股:修理:上海:資材
6504	富士電機(株)	制御:パワー:デバイス:電子:自動:香港:拠点:国内外:アジア:流通
6506	(株)安川電機	ロボット:韓国:スマート:半導体:従来:堅調:ユーロ:上海:背景:売
6674	(株)ジーエス・ユアサ コーポレーション	電池:用途:部:コーポレーション:機器:車:自動車:ユアサ:ジーエス:G S
6701	N E C	プラットフォーム:インフラ:行動:東京:中国:情報:開発:アメリカ:サービス:システム
6702	富士通(株)	携帯:ソフトウェア:効果:イノベーション:デジタル:排出:川崎:比率:テクノロジー:認識
6703	O K I	役務:プリンター:データ:連合:東京:システム:サービス:情報:製造:工事
6724	セイコーエプソン(株)	.※:液晶:強み:プリンター:ロボット:独自:オフィス:挑戦:転換:ユーロ
6752	パナソニック(株)	連合:蘇州:ネット:集中:全日本:上海:人材:ディスプレイ:指導:台湾
6758	ソニー(株)	マーケティング:兆:損害:量:事務所:中国:製造:東京:保険:エレクトロニクス
6762	T D K(株)	フィルム:製造:部品:投融資:開発:応用:変動:磁気:T D K:センサ
6770	アルプスアルパイン(株)	.※:スマート:ドイツ:旅行:上海:中国:独自:挑戦:業界:欧州
6773	パイオニア(株)	エレクトロニクス:運転:自動:現地:自動車:中国:製造:米国:情報:サービス
6841	横河電機(株)	機会:韓国:コーポレートガバナンス:プロセス:注力:姿:累計:上海:解決:挑戦
6857	(株)アドバンテスト	半導体:担当:試験:中国:製造:部品:情報:研究:開発:サービス
6902	(株)デンソー	エンジン:農業:愛知:車両:中華人民共和国:人材:電装:自動車:モノ:米国
6952	カシオ計算機(株)	渋谷:ユーロ:中国:製造:情報:本部:開発:東京:システム:サービス
6954	ファナック(株)	ロボット:製造:開発:システム:サービス:C N C:レーザ:ファナック:ロボドリル:サーボ
6971	京セラ(株)	液晶:ディスプレイ:複合:プリンター:原材料:半導体:素材:医療:様々:製造
6976	太陽誘電(株)	完成:韓国:原材料:責任:中国:製造:部品:国内外:機器:情報
6988	日東電工(株)	ライフ:人民元:次世代:台湾:韓国:上海:トップ:中国:製造:情報
7003	(株)三井E & Sホールディングス	海洋:船舶:三井:再生:請負:パイオ:転籍:プラント:電力:千葉
7004	日立造船(株)	集中:連合:発揮:ポートフォリオ:インフラ:機械:資源:中国:運転:処理
7011	三菱重工業(株)	宇宙:航空:航空機:エンジン:三菱:火力:健全:人材:西:防衛
7012	川崎重工業(株)	宇宙:エンジン:精密:蘇州:航空:船舶:輪:中華人民共和国:独自:ロボット
7013	(株)I H I	宇宙:ニューヨーク:江東:航空:商号:表示:修理:集中:連合:エンジン
7186	(株)コンコルディア・フィナンシャルグループ	金利:代行:各社:西:銀行:書き:圏:最終:金融:背景
7201	日産自動車(株)	連合:融資:車両:メキシコ:中国:自動車:防止:製造:東京:金融
7202	いすゞ自動車(株)	エンジン:車両:バス:中国:自動車:東京:製造:サービス:車:部品
7203	トヨタ自動車(株)	愛知:名古屋:ファイナンス:設立:運転:米国:中国:自動車:別:タイ
7205	日野自動車(株)	バス:インドネシア:タイ:製造:サポート:自動車:開発:アメリカ:車:部品
7211	三菱自動車(株)	インク:融資:カンパニー:ユーロ:自動車:インドネシア:米国:フィリピン:タイ:製造
7261	マツダ(株)	南:台:貸与:広島:自動車:次世代:中国:米国:製造:公表
7267	ホンダ	輪:ブラジル:埼玉:教育:カナダ:資源:尊重:カンパニー:米国:中国
7269	スズキ(株)	輪:台:.※:船:ユーロ:未来:車:中国:製造:開発
7270	(株)S U B A R U	宇宙:航空:連合:検査:完成:インク:不適切:調査:外部:ドル

表 4.4 頻出名詞リストのうち相関係数上位 10 個の名詞(7272~9009)

番号	企業名	頻出名詞リストのうち相関係数上位10個の名詞
7272	ヤマハ発動機(株)	輪:エンジン:プラットフォーム:静岡:ロボット:船:人材:ユーロ:増益:機械
7731	(株)ニコン	半導体:製造:機械:輸入:開発:サービス:東京:ヘルスケア:機器:品川
7733	オリンパス(株)	検査:痢(輸入:医療:ユーロ:東京:情報:兼務:開発:中国
7735	(株)SCREENホールディングス	半導体:指導:京都:台湾:印刷:ユーロ:代行:ドル:中国:製造
7751	キャノン(株)	貸与:複合:ディスプレイ:プリンター:1月:中国:中華人民共和国:兆:堅調:米国
7752	(株)リコー	複合:プリンター:損失:米州:譲渡:挑戦:オフィス:印刷:様々:データ
7762	シチズン時計(株)	プリンター:ハンド:不適切:精密:一時:フィリピン:無償:割当て:手続:中国
7911	凸版印刷(株)	デジタル:売:コミュニケーション:本社:印刷:回:ルール:情報:建:製造
7912	大日本印刷(株)	企画:安心:料:強み:さまざま:マテリアル:包装:責任:類:埼玉
7951	ヤマハ(株)	共和:完成:還元:ブランド:東京:開発:製造:機器:部品:電子
8001	伊藤忠商事(株)	輸出入:貿易:航空機:鉄鋼:融資:資材:石:天然:金属:非鉄
8002	丸紅(株)	輸出入:鉄鋼:ファイナンス:資材:融資:航空機:天然:石油:電力:非鉄
8015	豊田通商(株)	輸出入:融資:化学:金属:名古屋:現地:プロジェクト:中国:機械:自動車
8028	ユニー・ファミリーマートホールディングス(株)	沖縄:一方:家:最大限:小売:費用:転換:2月:カード:減
8031	三井物産(株)	天然:貿易:鉄鋼:石:次世代:現地:電力:ロンドン:金属:ブラジル
8035	東京エレクトロン(株)	半導体:期待:付加:サポート:次世代:データ:製造:痢(開発:サービス
8053	住友商事(株)	貿易:融資:現地:ファイナンス:金属:住友:ブラジル:資源:インフラ:人材
8058	三菱商事(株)	金属:現地:インフラ:石油:資源:三菱:化学:リース:東京:機械
8233	(株)高島屋	下:会員:昨年:ごと:獲得:クレジット:商業:百貨店:エリア:専門
8252	(株)丸井グループ	皆さま:テーマ:小売:創:都市:重点:モノ:商業:最終:カード
8253	(株)クレディセゾン	さまざま:ネット:ファイナンス:各社:提携:リース:不動産:データ:金融:東京
8267	イオン(株)	出店:皆さま:小売:健康:伸長:食:時間:千葉:発行:増減
8303	(株)新生銀行	預金:金利:貸借:金銭:各社:期待:機関:ファイナンス:集中:外部
8304	(株)あおぞら銀行	金利:預金:金銭:貸借:信託:機関:自己:銀行:集中:柱
8306	(株)三菱UFJフィナンシャル・グループ	金利:信託:各社:期待:高度:リテール:個人:政策:三菱:ニューヨーク
8308	(株)りそなホールディングス	信用:信託:高度:リテール:銀行:埼玉:機会:徹底:還元:配当
8309	三井住友トラスト・ホールディングス(株)	信託:金利:機関:自己:住友:政策:個人:高度:銀行:規制
8316	(株)三井住友フィナンシャルグループ	金利:リテール:高度:機関:住友:ファイナンス:個人:提携:三井:リース
8331	(株)千葉銀行	預金:信用:金銭:貸借:代行:採用:銀行:自己:書き:リース
8354	(株)ふくおかフィナンシャルグループ	福岡:フィナンシャルグループ:債権:書き:進化:銀行:ブランド:金融:サービス:情報
8355	(株)静岡銀行	預金:信用:金銭:機関:提携:リース:銀行:健全:コア:静岡
8411	(株)みずほフィナンシャルグループ	金利:信託:機関:リテール:高度:政策:個人:自己:規制:徹底
8601	(株)大和証券グループ本社	貸借:江東:東京:情報:金融:サービス:本社:銀行:運用:千代田
8604	野村ホールディングス(株)	信託:貸借:金銭:ニューヨーク:ロンドン:別:シンガポール:調査:付加:東京
8628	松井証券(株)	信託:信用:拡充:個人:家:業界:サービス:システム:オンライン:手数料
8630	SOMPOホールディングス(株)	修正:生命:損害:締結:保険:責任:安心:ロンドン:ジャパン:シンガポール
8725	MS & ADインシュアランスグループホールディングス(株)	修正:締結:生命:たち:住友:損害:保険:安心:三井:金融
8729	ソニーフィナンシャルホールディングス(株)	修正:生命:各社:たち:私:設立:保険:損害:個人:銀行
8750	第一生命ホールディングス(株)	生命:修正:各社:たち:私:発揮:期待:保険:締結:分の
8766	東京海上ホールディングス(株)	修正:生命:発揮:締結:損害:保険:シンガポール:一般:東京:金融
8795	(株)T & Dホールディングス	生命:保険:政策:締結:代行:東京:サービス:運用:情報:フィナンシャル
8801	三井不動産(株)	分譲:オフィス:住宅:事務所:ホテル:元:新宿:不動産:商業:千葉
8802	三菱地所(株)	分譲:マンション:戸:オフィス:住宅:ホテル:ネット:強み:場:不動産
8804	東京建物(株)	分譲:マンション:戸:住宅:オフィス:事務所:不動産:場:計:通り
8830	住友不動産(株)	分譲:マンション:住宅:事務所:不動産:兆:新宿:ルール:遵守:定時
9001	東武鉄道(株)	沿線:ホテル:バス:運輸:百貨店:安心:不動産:無償:割当て:勧告
9005	東京急行電鉄(株)	沿線:ホテル:バス:渋谷:園:都市:百貨店:不動産:住宅:健康
9007	小田急電鉄(株)	沿線:駅:ホテル:バス:百貨店:運輸:渋谷:対策:テーマ:安心
9008	京王電鉄(株)	沿線:駅:ホテル:百貨店:バス:運輸:渋谷:対策:新宿:不動産
9009	京成電鉄(株)	沿線:ホテル:バス:運輸:百貨店:安心:車両:千葉:不動産:選任

表 4.5 頻出名詞リストのうち相関係数上位 10 個の名詞(9020~9984)

番号	企業名	頻出名詞リストのうち相関係数上位10個の名詞
9020	東日本旅客鉄道(株)	駅:全日本:ホテル:渋谷:運輸:小売:圏:飲食:車両:対策
9021	西日本旅客鉄道(株)	駅:沿線:全日本:ホテル:用地:都市:運輸:小売:飲食:百貨店
9022	東海旅客鉄道(株)	駅:全日本:用地:運輸:対策:不動産:静岡:本部:線:車両
9062	日本通運(株)	都市:オセアニア:南:航空:サポート:運送:貨物:エリア:重点:最終
9064	ヤマトホールディングス(株)	優先:沖縄:満足:方:運輸:物流:東京:情報:サービス:開発
9101	日本郵船(株)	船:貨物:海洋:燃料:船舶:運送:航空:横浜:名古屋:-(株)
9104	(株)商船三井	海洋:船:貨物:船舶:燃料:神戸:三井:統合:横浜:計
9107	川崎汽船(株)	船:海洋:貨物:燃料:運送:船舶:ポートフォリオ:転換:統合:量
9202	A N Aホールディングス(株)	全日本:圏:千葉:旅行:沖縄:-(株):航空:貨物:航空機:ポートフォリオ
9301	三菱倉庫(株)	神戸:運輸:運送:不動産:貨物:インドネシア:大阪:東京:上海:情報
9412	(株)スカパー J S A Tホールディングス	体:従来:地球:拡充:宇宙:インターネット:獲得:プラットフォーム:東京:運用
9432	日本電信電話(株)	移動:パートナー:データ:研究:東京:開発:情報:サービス:北米:システム
9433	K D D I (株)	移動:役務:とき:合併:規定:範囲:いずれ:適正:違反:沖縄
9437	(株)N T Tドコモ	移動:役務:とき:合併:規定:パートナー:範囲:適正:いずれ:違反
9501	東京電力ホールディングス(株)	火力:用地:電:-(株):江東:都市:パートナー:調査:対策:財
9502	中部電力(株)	火力:電:線:天然:名古屋:小売:範囲:電力:運転:不動産
9503	関西電力(株)	電:柱:トップ:西:北:程度:請負:調査:重点:電力
9531	東京ガス(株)	皆さま:都市:天然:転籍:不動産:リース:電力:オーストラリア:(株):東京
9532	大阪ガス(株)	都市:天然:電力:西:不動産:リース:熱:スポーツ:オーストラリア:転籍
9602	東宝(株)	飲食:企画:減:前年度:税:固定:2月:本:不動産:収入
9613	(株)N T Tデータ	公共:付加:データ:江東:本部:ユーロ:研究:ブランド:情報:中国
9681	(株)東京ドーム	飲食:同等:水準:スポーツ:ホテル:役:減:不動産:損失:1月
9735	セコム(株)	渋谷:プラットフォーム:医療:安心:さまざま:指導:不動産:損害:データ:メディカル
9766	コナミホールディングス(株)	スポーツ:時間:健康:規制:米:国:情報:東京:資源:サービス:システム
9983	(株)ファーストリテイリング	出店:ブランド:米:国:中国:プロジェクト:上海:開発:情報:ユニ:クロ
9984	ソフトバンクグループ(株)	合併:移動:号:規制:作成:以外:負債:ベース:固定:分の

参考文献

- [1] Bollen, J., Mao, H. and Zeng, X.-J. 2010. Twitter mood predicts the stock market. Journal of Computational Science 2(1)
- [2] 佐藤史仁、佐久間洋明、小寺俊哉：テキストマイニングによる有価証券報告書の因果関係文の抽出、日興リサーチレビュー、2017年10月号（2017）
- [3] 和泉潔、後藤卓、松井藤五郎：経済テキスト情報を用いた長期的な市場動向推定、情報処理学会誌、Vol.52、No.12、pp.3309-3315（2011）
- [4] 会社四季報 2018年4集（秋号）
- [5] Wilson, T, Hoffmann, P, Somasundaran, S, Kessler, J, Wieve, J, Choi, Y, Cardie, C, Riloff, E, & Patwardhan, S. 2005. OpinionFinder: A system for subjectivity analysis. pp. 34-35