

令和 5 年 6 月 22 日現在

機関番号：34506

研究種目：基盤研究(B)（一般）

研究期間：2020～2022

課題番号：20H01561

研究課題名（和文）テキストマイニングの手法を活用した有価証券報告書の実態分析および実証分析

研究課題名（英文）Empirical Study of Annual Report using Text Mining Techniques

研究代表者

伊藤 健顕 (Ito, Takeaki)

甲南大学・マネジメント創造学部・准教授

研究者番号：00709496

交付決定額（研究期間全体）：（直接経費） 5,400,000円

研究成果の概要（和文）：本研究の目的は、企業の開示している記述情報に対して、テキストマイニングのアプローチを用いて分析することでその開示実態を明らかにし、さらに価値関連性や経済的帰結といった有用性を検討することである。具体的には日本企業が開示している有価証券報告書に記載されるテキスト情報を分析対象とし分析を行った。分析の結果、テキストマイニングの技術を活用した長期的かつ包括的な分析、決算短信におけるコロナウイルス感染症に関連した情報開示、経営者交代と記述情報の関連性、有価証券報告書の記述情報のトーンと読性が中長期的な将来業績に与える影響について明らかにすることができた。

研究成果の学術的意義や社会的意義

本研究で行ったような有価証券報告書を対象とした長期間での分析はこれまで存在しなかった。またそれらのデータを活用して経営者交代や業績予想といった企業の特徴との関連を分析したことにより、情報開示と企業行動に関する知見を蓄積することができたと考えられる。これらの研究結果から情報作成者である企業と情報利用者である投資家やその他のステークホルダーとの情報の非対称性を最小化するためのヒントを得ることができる可能性があると考えられる。

研究成果の概要（英文）：The purpose of this study is to utilize text mining techniques to analyze the descriptive information disclosed in the financial statements of Japanese companies, in order to gain insights into the actual state of disclosure and examine its value relevance and economic consequences. Specifically, we conducted an analysis of the textual information found in the annual reports released by Japanese companies. Through this analysis, we were able to perform a comprehensive and long-term assessment using text mining techniques. Furthermore, we investigated the relationship between the disclosure of information regarding coronavirus infection in financial statements, management changes, and descriptive information, as well as the impact of tone and readability of descriptive information in annual reports on the medium- to long-term future business performance.

研究分野：会計学

キーワード：テキストマイニング ディスクロージャー 有価証券報告書 実証分析

1. 研究開始当初の背景

本研究課題の核心をなす学術的「問い」は「有価証券報告書に記載されている記述情報はステイクホルダーにたいして有用な情報を提供しているのか」である。

その背景には、財務会計およびファイナンス分野における実証研究アプローチの多様化が挙げられる。

情報には定量情報だけでなく、文章による記述情報も多分に含まれている。そして記述情報を定量化し実証研究を行うというテキストマイニングの手法を活用したアプローチは特に米国を中心とした海外では先行研究が存在し、研究成果が蓄積されてきている。

テキストマイニングのアプローチを用いた研究は大きく2つに分けることができる。1つは可読性(readability)に関する研究であり、いまひとつは内容分析(content analysis)に関する研究である。

可読性に関する研究は可読性の指標(Fog index など)の有用性を検証したもののや可読性の指標と財務指標との関連を検証したものが中心である。たとえば Li[2008]では業績の悪い企業ほどアニュアルレポートの可読性が低く文章が長くなる傾向があり、業績がよい企業ほど可読性が高くなる(読みやすい)という傾向があることが示されている。

また、内容分析に関する研究の中心はトーン(tone)に関する研究である。トーンとは文書に記載されている単語ひとつひとつをポジティブなものとネガティブなものに分類し、その文書全体におけるポジティブな単語とネガティブな単語の割合をもって、ポジティブなトーンおよびネガティブなトーンと判断し分析を行うものである。トーンは企業の業績や財務プロファイルなど様々な企業特性と関連を示していることが先行研究においても示されている(Loughran and McDonald[2016]など)。

これらの研究は米国を中心に研究が蓄積されてきているが、日本においては十分に研究が行われているとは言い難い。そのような状況の中、2019年3月には金融庁より有価証券報告書における記述情報の整理を目的とした「記述情報の開示に関する原則」が公表された。具体的には「経営方針・経営戦略等」「経営成績等の分析(MD&A)」「リスク情報」の3点を中心に望ましい開示に向けた取組みについての原則が示されている。「記述情報の開示に関する原則」が公表された経緯としては、金融審議会におけるディスクロージャーワーキンググループにおいて有価証券報告書における情報開示が不十分であると指摘されたことが挙げられる(たとえば MD&A については大谷・上利・堀内・岡村 [2018]を参照)。これらの指摘は実務家を中心とした意見であるが、研究者による記述情報の開示傾向や開示による経済効果についての分析は数が少ない。例えば、首藤・緒方(2009)は2006年度において上場企業300社のMD&A情報を対象とした分析を行っている。分析の結果、新規上場企業がMD&Aの開示に積極的であること、開示項目については多少の偏りがあること、MD&Aの記述において他の箇所の参照を指示している企業が少なくないこと、が明らかにされている。

また、中野(2010)はリスク情報とMD&A情報を分析対象とし、その開示量と開示内容を分析している。分析期間は2004年3月期から2007年3月期までとし、東証第一部上場企業を対象に分析を行っている。その結果MD&Aについて、開示量は2005年以降、平均的な開示量は増えておらず、記載すべき内容はすでに固まっている傾向がある、業種別にみると、開示に積極的な企業とそうではない業種が存在する、事業リスクの高い企業、大規模企業、市場からの注目度の高い企業および事業構造が複雑な企業ほど、積極的に開示する傾向がある、という結果を報告している。開示内容については、全般的に、同様の表記に従い記述しており、標準化されているといつてよい。表題は標準化されているものの、その具体的記述内容に差異があると推測される、という結果が報告されている。

以上のように、実務においては記述情報の問題点や課題についての指摘が存在するものの、研究者によってMD&A情報をはじめとした記述情報の有用性や効果が十分に検討されているとは言い難い。また実証分析以前の問題として、記述情報の場合は会計情報や財務情報のような定量情報とは異なり値の分布を要約することは容易ではない。しかし、テキストマイニングやそれに関連した様々な手法の発展により、そのような性格をもつ記述情報についても傾向を定量的に把握することが可能となってきている。

また、分析対象である有価証券報告書の有用性については検討の余地が存在する。適時性の視点からは有価証券報告書に記載される情報にどれだけの有用性があるかは明確にされていない。財務諸表を中心とした決算情報は決算速報や決算短信において先んじて情報開示されるため、株主総会での承認後に公開される有価証券報告書にどれだけの情報内容があるかは財務諸表情報以外の情報内容を検討する必要がある。

2. 研究の目的

本研究の目的は、企業の開示している記述情報に対して、テキストマイニングのアプローチを

用いて分析することでその開示実態を明らかにし、さらに価値関連性や経済的帰結といった有用性を検討することである。

これまでの財務会計・ファイナンス分野では会計数値に代表される定量情報を用いた実証分析が中心であった。しかし、実際に企業の開示している情報には定量情報だけではなく、文章による記述情報も多分に含まれている。そこで、本研究ではそれらに含まれる文章の情報がどのような役割を持ち、影響を与えているかをテキストマイニングの手法を用いて定量化し、実証分析のモデルへと組み込む。日本語を対象としたテキストマイニングは英語と比較した場合にはその言語的特性から難易度が高く、分析が困難であった。しかし、近年では R や Python といったプログラミング言語において日本語を取り扱うことも比較的容易になり、また、PC の性能が高まったこともあり日本語のテキストマイニングを行うことも可能である。それらの技術を用いて定性情報をディスクロージャー研究の分析モデルに組み込むことにより、これまでの定量情報のみを用いた実証研究では明らかになっていない事実を発見できる可能性がある。

Lev and Gu[2016]が指摘するように、会計情報の有用性は時系列で低下する傾向にある。そこで会計情報以外の情報として記述情報に注目し、分析することで会計情報の有用性低下という問題に対して新たな知見を得ることができると考えられる。そして、上述のようにこれまでの定量情報のみで分析されてきた分野の中でも記述情報が重要な役割を占めている可能性がある場合には、定量情報のみを用いた分析だけではなく記述情報の分析も合わせて行うことで新たな知見を得ることができると考えられる。

具体的には、記述情報は財務情報を補完する役割が期待されている。したがって、会計情報を所与としたうえで記述情報に増分情報内容があるかどうかを明らかにする必要がある。その増分情報内容を確認するためにも記述情報をテキストマイニング等の技術を活用して定量化し、分析モデルへと組み込む必要がある。

また、本研究では日本企業の記述情報と米国を中心とした海外企業の記述情報との比較分析も視野に入れている。米国企業を対象としたテキストマイニング研究は蓄積されつつあるが、言語の異なる他国企業との比較を行った研究は筆者が調べたところ存在しない。この点については大きな貢献となりうる。

3. 研究の方法

上述のように、本研究では記述情報として具体的には有価証券報告書を分析対象とし、その内容について詳細に検討する。有価証券報告書の中でも「事業の状況」および「提出会社の状況」を中心に分析を行う。「事業の状況」の中でも特に「事業等のリスク」「財政状態、経営成績及びキャッシュ・フローの状況の分析」を中心に分析を行う。

分析視点としては第一に有価証券報告書における記述情報の開示実態を明らかにすることである。本研究の研究対象の中心である有価証券報告書における記述情報の中でも特に「事業等のリスク」「MD&A」「コーポレート・ガバナンスの状況」については2004年3月に始まる事業年度より開示が義務付けられた。その後、様々に議論がされ、近年では「記述情報の開示に関する原則」が公表されているがその開示実態については必ずしも明らかになっていない。そこで本研究では有価証券報告書における記述情報の開示がどのように変化しているのかを明らかにする。

第二の分析視点は記述情報と定量情報の関連を明らかにすることである。海外での先行研究では記述情報と業績の関連や株価の関連について様々に議論がされている。日本においても先行研究は存在するものの、十分に検討されているとは言い難い。

具体的には記述情報のトーンや可読性といった視点や、機械学習等の手法を用いてより詳細に分析を行い、企業の業績や株価、企業価値などどのような関連があるのかを検証する。複数の指標を活用し分析することで、有価証券報告書の有用性について検討を行う。

4. 研究成果

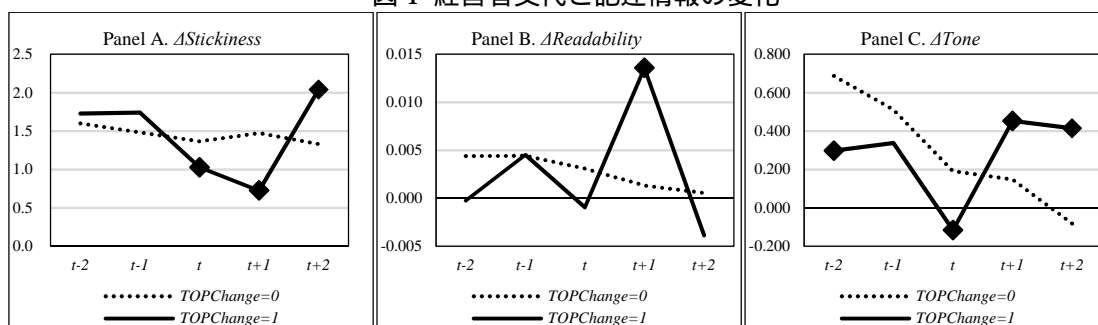
2020年度は有価証券報告書に記載される記述情報について、本研究で用いる記述情報に関するデータベースの構築、テキストマイニングの技術を活用した長期的かつ包括的な分析、決算短信における新型コロナウイルス感染症に関連した情報開示に注目し分析を行った。の結果、有価証券報告書の記述情報は全体的に増加傾向にあり、近年では特にMD&Aとガバナンス情報において顕著な増加傾向が見て取れた。また、MD&Aやリスク情報は業績や財政状況がテキスト情報の量やその質と関係していることがわかった。の結果、共起ネットワークの分析から、多くの企業にとって決算短信の中心的な内容が「COVID-19の感染拡大が売上高と資金、および財政状態にどのような影響を与えたのか」であったことが明らかになった。さらに、センチメント分析から、売上高が減少した企業そして次期の業績予想を開示しなかった企業ほど新型コロナウイルス感染症に関連した単語や不確実性に関連した単語を多く使いながら相対的にネガティブに説明していること、決算短信の開示が遅れた企業ほど新型コロナウイルス感染症に関連した単語を多く使っておりその説明もネガティブであること、一方で事前にBCP(事業継続計画)を開示した企業ほど不確実性に関連した単語の使用が多く、かつその説明がポジティブであることが観察された。

2021年度は有価証券報告書に記載される記述情報について、テキストマイニングの技術を

活用した長期的かつ包括的な分析、経営者交代と記述情報の関連性に注目し分析を行った。の結果は以下の通りである。有価証券報告書の記述情報は全体的に増加傾向にあり、近年では特に MD&A とガバナンス情報において顕著な増加傾向が見取れた。また、MD&A やリスク情報は業績や財政状況がテキスト情報の量やその質と関係していることがわかった。特に、財政状況や業績の悪い企業ほど開示量が多く、業績の良い企業の可読性が高いとの結果は先行研究とも整合的である。さらに、業績の良い企業ほど MD&A の記述情報がポジティブになっていた。また、企業の規模は MD&A やリスク情報だけでなくガバナンス情報とも関係していることがわかった。その他、監査を受ける監査法人の大きさや企業の上場マーケット、そして会計基準なども企業の記述情報と関連していた。また、の結果、経営者交代によって記述情報のスティッキネス(過年度の記述情報の再利用度合い)が低下するとともにその可読性が向上すること、さらに記述情報のトーンがポジティブになることが明らかになった。

2022 年度は有価証券報告書に記載される記述情報について、経営者交代と記述情報の関連性、有価証券報告書の記述情報のトーンと可読性が中長期的な将来業績に与える影響、に注目し分析を行った。の結果、経営者交代によって記述情報のスティッキネス(過年度の記述情報の再利用度合い)が低下するとともにその可読性が向上すること、さらに記述情報のトーンがポジティブになることが示された(矢澤・金・伊藤[2022]、図 1)。

図 1 経営者交代と記述情報の変化



このような変化は、有価証券報告書の提出まで十分な時間があり、新任の経営者がその作成手続に実質的な影響を及ぼすことが可能であると考えられる、交代後第 2 期目に提出される記述情報において顕著に観察された。さらに、経営者交代による記述情報の変化は新任経営者の属性や記述情報の記載されるセクションによって異なることも明らかにされた。このような発見事項は、経営者が有価証券報告書における記述情報を通じて投資家とコミュニケーションを行っていることの証拠といえる。また、の結果は以下の通りである。有価証券報告書に記載される記述情報の中でも MD&A には従来観察されていたよりもより長期的な業績に関する経営者の見通しが反映されており、また、対処すべき課題にも将来業績に関する経営者の見通しが反映されていることが示唆された。これらの知見は、中長期的な企業価値の向上に向けた情報作成者と利用者の間の対話に資するものであろう。

引用文献

1. Lev, B., and F.Gu. (2016). *The End of Accounting and the Path Forward for Investors and Managers*. Wiley.
2. Li, F.(2008). Annual Report Readability, Current Earnings, and Earnings Persistence. *Journal of Accounting and Economics*, 45, 221-247.
3. Loughran, T., and McDonald, B. (2014). Measuring Readability in Financial Disclosures. *Journal of Finance*, 69(4), 1643-1671.
4. 大谷潤・上利悟史・堀内隼・岡村健史「企業内容等の開示に関する内閣府令等改正の解説 非財務情報の開示充実等」『週刊商事法務』, 第 2163 号, 2018 年 4 月.
5. 金融庁「記述情報の開示に関する原則」2019 年 3 月.
6. 首藤昭信・緒方英明(2009)「実務研究 有価証券報告書における「財政状態及び経営成績の分析(MD&A)」について」『研究所レポート』プロネクサス総合研究所, 第 3 号.
7. 中野貴之(2010)「財務諸表外情報の開示実態 事業等のリスクおよび MD&A の分析」『財務 諸表外情報の開示と保証 ナラティブ・リポーティングの保証』同文館出版, 2010 年 10 月.
8. 矢澤憲一・金鉉玉・伊藤健顕「経営者交代が記述情報の変化に与える影響」『会計プログレス』第 23 号, 2022 年 9 月.

5. 主な発表論文等

〔雑誌論文〕 計7件（うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 0件）

1. 著者名 矢澤憲一, 金鉉玉, 伊藤健顕	4. 巻 60
2. 論文標題 会計研究の新展開シリーズ (第 2 回) 財務報告研究におけるテキスト分析	5. 発行年 2022年
3. 雑誌名 証券アナリストジャーナル	6. 最初と最後の頁 52 - 61
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 金鉉玉, 矢澤憲一, 伊藤健顕	4. 巻 2022
2. 論文標題 経営者交代が記述情報の変化に与える影響	5. 発行年 2022年
3. 雑誌名 会計プロGRESS	6. 最初と最後の頁 49 - 67
掲載論文のDOI (デジタルオブジェクト識別子) 10.34605/jaa.2022.23_49	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 廣瀬喜貴	4. 巻 60
2. 論文標題 有価証券報告書における定性情報が投資家の意思決定に与える影響 : オンラインサーベイ実験による検証;特集 有価証券報告書の定性情報	5. 発行年 2022年
3. 雑誌名 証券アナリストジャーナル	6. 最初と最後の頁 18 - 26
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 矢澤憲一, 金鉉玉, 伊藤健顕	4. 巻 74
2. 論文標題 有報の60年	5. 発行年 2022年
3. 雑誌名 企業会計	6. 最初と最後の頁 27-37
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 矢澤憲一, 伊藤健頭, 金鉉玉	4. 巻 56
2. 論文標題 テキストマイニングを用いたMD&A、リスク、ガバナンス情報の分析	5. 発行年 2021年
3. 雑誌名 青山経営論集	6. 最初と最後の頁 60 - 84
掲載論文のDOI (デジタルオブジェクト識別子) 10.34321/22038	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 矢澤憲一, 伊藤健頭, 金鉉玉	4. 巻 72
2. 論文標題 テキストマイニングでみてきた日本企業開示の特徴 2020年3月期決算短信におけるコロナディスクロージャーの分析	5. 発行年 2020年
3. 雑誌名 企業会計	6. 最初と最後の頁 84 - 92
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 矢澤憲一, 伊藤健頭, 金鉉玉	4. 巻 72
2. 論文標題 テキストマイニングでみてきた日本企業開示の特徴 全上場企業15年分・のべ55,000社の有報記述情報分析	5. 発行年 2020年
3. 雑誌名 企業会計	6. 最初と最後の頁 105 - 113
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計6件 (うち招待講演 0件 / うち国際学会 0件)

1. 発表者名 廣瀬喜貴, 伊藤健頭
2. 発表標題 ディスクロージャーのトーンとリーダービリティが情報利用者に与える影響
3. 学会等名 行動経済学会第16回大会
4. 発表年 2022年

1. 発表者名 廣瀬喜貴
2. 発表標題 開示情報のリーダビリティと個人投資家の意思決定の関連: オンライン・サーベイ実験による検証
3. 学会等名 日本会計研究学会第81回全国大会
4. 発表年 2022年

1. 発表者名 矢澤憲一, 金鉉玉, 伊藤健顕
2. 発表標題 有価証券報告書は何を語るのか - テキスト解析で辿る60年間の変遷
3. 学会等名 日本IR学会
4. 発表年 2021年

1. 発表者名 矢澤憲一, 伊藤健顕, 金鉉玉
2. 発表標題 経営者交代が有価証券報告書の記述情報に与える影響について
3. 学会等名 日本会計研究学会第80回大会
4. 発表年 2021年

1. 発表者名 矢澤憲一, 伊藤健顕, 金鉉玉
2. 発表標題 テキストマイニングを用いた我が国有価証券報告書の記述情報の分析
3. 学会等名 日本会計研究学会第79回大会
4. 発表年 2020年

1. 発表者名 矢澤憲一, 伊藤健頭, 金鉉玉
2. 発表標題 テキストマイニングを用いた我が国国有価証券報告書の記述情報の分析
3. 学会等名 日本経済会計学会第37回年次大会
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	金 鉉玉 (Kim Hyonok) (40547270)	東京経済大学・経営学部・教授 (32649)	
研究分担者	佐藤 圭 (Sato Kei) (80826602)	甲南大学・経営学部・准教授 (34506)	
研究分担者	廣瀬 喜貴 (Hirose Yoshitaka) (90760265)	大阪公立大学・大学院経営学研究科・准教授 (24405)	
研究分担者	矢澤 憲一 (Yazawa Kenichi) (70406817)	青山学院大学・経営学部・教授 (32601)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------