

令和元年6月21日現在

機関番号：34425

研究種目：基盤研究(C)（一般）

研究期間：2015～2018

課題番号：15K03802

研究課題名（和文）テキストマイニングによる環境報告書の記述的側面の分析

研究課題名（英文）Descriptive analysis of environmental reports using text mining technique

研究代表者

前田 利之（Maeda, Toshiyuki）

阪南大学・経営情報学部・教授

研究者番号：70320041

交付決定額（研究期間全体）：（直接経費） 2,900,000円

研究成果の概要（和文）：本研究では、環境報告書に記載されたトップメッセージに対してテキストマイニングを行い、環境保全コスト率の高さと出現頻度するキーワードとの関係を視覚的に検証した。分析結果から、環境コストを多く支出する企業は、社会とともにCSRや環境問題に対して重大な関心を持ち、新製品の開発や環境対応製品の事業化など積極的な活動を行っており、支出が少ない企業群は、環境関連活動に具体的な目標を持たず、経営改善の一環としてCSR活動を行っている傾向にあることが観察された。

研究成果の学術的意義や社会的意義

近年の環境対応製品の事業化により、企業の「環境への対応」の動機と目的は更に複雑化しているといえ、この研究においては企業や産業の知識だけではなく、社会や地域といったより広範な知識の重要性が増してきているといえる。トップメッセージのテキスト分析を環境コストの金額という定量的な情報と結び付けてやることで、環境に対する企業の姿勢を直観的な形で浮かび上がらせるという当初の研究目標は達成することができたといえる。

研究成果の概要（英文）：In this study, we introduce text mining technique to the top message described in the environmental report, and the relationship between the high environmental protection cost rate and the frequency of occurrence keywords was visually verified. Companies that spend a lot of environmental costs have a serious interest in CSR and environmental issues with society, and they are actively engaged in development of new products and commercialization of environment-friendly products, and the group of companies that spend less It has been observed that there is a tendency for CSR activities to be carried out as part of management improvement without specific targets for environment-related activities.

研究分野：データサイエンス

キーワード：テキストマイニング 環境報告書 時系列分析 トップメッセージ

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

環境に対する企業の取り組みには、利害関係者から多大な関心が寄せられている。環境改善活動に関する成果情報が必要とされる背景には、かかる情報に企業レベルでの環境汚染を抑制するチェック機能が期待されているからである。それだけでなく、環境規制に抵触する可能性が高まれば、社会的信用だけでなく経済的な評価をも損なうことになるから、そうした可能性を事前に予測することが求められる。Berthelot et al. (2003) が主張するように、環境改善に関わる姿勢は、企業の持続可能性を見極める上で重要な指針を与えるのである。

ただし、企業が提供する環境情報は、膨大な記述情報から構成されることが多い。Neu et al. (1998) によれば、意図しているかどうかに関わらず、経営者は語彙や表現を工夫することで、企業に対する利害関係者の印象に働きかける傾向がある。たしかに、積極的な見出しのもとで自身の行動を正当化する記述が多用されるとき、情報の受け手は企業イメージを上方に改定する可能性が高い。むしろ、慎重な記述で将来ありうべきリスクを喚起するような表現は、企業の姿勢について否定的な連想を生じさせかねない。他方、環境会計情報の開示が実務に浸透した結果、環境改善に向けた活動目標に内実が備わるようになった。もちろん、制度で強制されない環境会計情報を任意に開示すること自体、社会的責任の遂行に対する自信の表白かもしれない。しかし、本業が不調な場合など、環境コストの負担を絞らざるをえない局面もあろう。それに伴う企業イメージの毀損を、定性的な記述を駆使することで償うことは可能であろうか。このような問題意識に基づき本研究では、テキストマイニングを活用して、環境保全コストの大きさと環境情報の記述内容との関係を明らかにする。

2. 研究の目的

本研究では、テキストマイニングの技法を用いて、定量的なデータと定性的な記述との間の関連を分析することで、企業の行動の動機、とりわけ上場企業が取り組む環境関連活動がどのような経営理念や経営戦略によって促進されているかを、環境報告書から読み取ることを目的とする。近年、CSR (企業の社会的責任) の普及によって、企業による社会関連情報の公開は拡充が進んでおり、更に様々な研究においてテキストマイニングの技法が主要な情報収集の手段として普及している。こうした状況はテキストマイニングを用いた研究にとって追い風となっているが、マーケティング分野を除く経営研究において、この技法の活用はあまり進んでいない。本研究では、2000年度から2011年度の環境報告書に対するテキストマイニングを行うことで、積極性の規定要因の変遷を跡づけ、より詳細な分析結果を導き出すことを目的とする。

3. 研究の方法

3.1 解析対象テキストについて

本研究では東証一部上場企業であり、Web上に環境報告書を掲示している企業のうち、2000年度から2011年度までの12年間の環境報告を行った環境、社会貢献報告書の揃う企業の環境報告書のトップメッセージを分析対象とする。CSRの重要性の認識が広がる現代では、こうした報告書は経営責任者の誓約や環境経営の方針、そして現状の環境保全活動、将来の目標など自社の価値や意義を消費者やステークホルダーに対して発信する重要な舞台であり、人々が最初に目にするトップメッセージには、それらの総括となる内容が記載されると考えられる。

3.2 分析方法

研究の内容は上記の環境報告書のトップメッセージを抽出し、それらの報告書を企業単位で業種ごとに売上高に占める環境コスト比率で上位3社、下位3社でグループ分けを行い、時系列での変化の動向を捉えやすいよう1期4年毎の3期にさらに分割してこの12年間におけるトップメッセージの変化に対する分析を行う。

売上高に占める環境コスト比率で分析を行った理由は、企業規模に左右されない分析を行う事、業種ごとの環境コスト比率の特徴を見出すという2点の理由からである。

また、時系列の分析のため12年を4年毎3期に分割した理由は、まず期間ごとのグループの資料数が同数となるため比較が容易であること、時系列による変化の分析の際にグループ数が多くなりすぎないことがあげられる。

トップメッセージから抽出したデータは以下のようにグループ化と正規化を行う。

(1) 産業毎にコスト率が大きい順と小さい順に3社のグループに分類し、それぞれのグループを「上位」「下位」グループと呼ぶ。

(2) 上記のグループを報告書の発行年から更に2000~2003、2004~2007、2008~2011の期間で分け、それぞれを前期、中期、後期と呼ぶ。

(3) トップメッセージから抽出する形態素を名詞(固有名詞、人名、地名は除く)と形容詞のみに対象を限定する。

(4) 報告書ごとの単語使用頻度に対して200語で正規化を行い、単語使用頻度が大きい順に上位50語を抽出する。但し、「年」、「よう」、「億」など分析上意味をなさないキーワードは除く。

(5) 各グループを分析に使用する形で統合し、2つ以上のグループで使用されているキーワードを抽出する。

以上を行った上で、キーワードとそれぞれのグループにおける出現頻度の関係を視覚的に確認するために対応分析を行い、そこから論理的に導かれる考察を提示する。

また、多業種による解析を行う際の具体的な手順は上記と同様に行い、コスト比率や時期ごとを

基準としてグループをまとめた後、同様に対応分析を行いその結果を考察する。

4. 研究成果

4.1 業種別形態素単語出現頻度

建設の単語出現頻度では環境コスト比率が高く、事業が環境に関係の深い業種であるためか、比率の異なる他業種と比較して全期間を通して環境に関する語が多く見られた。他にもコスト比率の高い業種はコスト比率の低い業種と比較して具体的製品に多く言及する傾向にあるが、建設は社会に関するキーワードも多い。事業が利用者の生活に深く関係し、影響範囲が広いため社会全体への言及が多くなっていると考えられる。

化学の単語出現頻度では他のコスト比率の高い企業同様に製品や環境に関する語が多く見られるが、多くの場合それは減少傾向にある。それに反比例するように活動に関する語が増加傾向にあるため、CSR 技術開発が中期までに一定段階に到達したため、その他の活動を積極化させていると考えられる。

輸送の単語出現頻度では上位 30 語に特に製品に関する語が多く、「ディーゼルエンジン」と「トラック」といった名前に踏み込んだ数少ない例である。そのためか他業種と比較しても製品に関する語が非常に多く、それが環境コスト比率の高いグループ全体の結果にも大きな影響を与えていると言える。しかし、「ディーゼルエンジン」と「トラック」は 1 社が繰り返し発言した結果が強く表れているため、企業数が少ないために現れた影響であるといえる。また、事業が環境に関係の深い業種であるためか、比率の異なる他業種と比較して全期間を通して「環境」が多く使われている。

鉄非鉄単語頻出度表は環境関連キーワードの大幅な減少に伴い、多くの業種と同様に後期にかけて CSR 関連の語が増加しているが、景気悪化の影響が各キーワードの増加が少なく、活動が停滞している印象がある。

家電では上位 30 語の中では下位に位置するもののステークホルダー関連のキーワードが特徴的である。文化、やさしいなどのキーワードは一旦中期で途切れながらも後期で再度復活するなど、トップメッセージから顧客に対する CSR 以上の取り組みが察せられる。それらの影響で他業種と比較して環境関連のキーワードは少なくなっている。

食品業界では食品や調味料といった消費者の健康や生活に直接関係する業種であるため、他の業種とは違った傾向を見せている。環境問題を焦点としていた前期以降に「環境」に関連する語句の減少とともに、多くの企業は企業活動と社会貢献の語句が増加するが、社会と文化への貢献を軸とした単語が大幅に増加している。環境投資比率の変遷を考慮すると、環境報告書が CSR 報告書に転じた際に比較的環境への影響の少ない業種であったため、新たな企業活動ではなく、既存製品の品質や利用先についての言及が増加したと考えられる。

制御の最も特徴的な点として「環境」の語より「社会」の形態素の頻度が多いという類を見ない特徴がある。特にこの業種の「社会」の頻度は前期から一貫して多く、環境報告書のガイドライン制定以前から社会を見据えて行動してきたといえる。その分環境に関する言及が少なく、一部の時期を除いて環境コスト比率を一定に保っていたキーワードの頻度であるといえる。

通信では環境コスト比率が中程度の業種よりも環境に対する言及が多く見られるが、環境関連のキーワードの減少に伴う経営関連のキーワードの増加は目立つものの技術関連のキーワードがほとんど見られないため、経営関連のキーワードの増加は目的や将来像の言及にとどまり、景気悪化に伴って活動が消極的になっていると取れる。

精密の単語出現頻度では精密に医療機器関連業種が多く分類される影響から、消費者の生命や生活に直接関係する業種であるため、食品同様に他の業種とは違った傾向を見せている。環境問題を焦点としていた前期以降に「環境」に関連する語句の減少とともに社会関連、特に医療や利用者に関連する単語が大幅に増加している。食品同様に業種内の環境問題が小さく、それ以上に重大な危機に直面する利用者への配慮が表れた形と考えられる。

4.2 対応分析

この節ではキーワードを選別し対応分析を行い、その結果に対する考察を行う。

建設の対応分析では左側に前期、右側に中期後期が配置され、一部で期間ごとの繋がりを無視したグループを形成していることが特徴的である。キーワードの配置から推測すると上位のグループは前期に負荷低減や廃棄物削減の技術の近くに配置されているため、中期にそれらを利用して積極的な活動に取り組んでいたものの、東日本大震災や景気の下落の影響からその活動を見直し、より CSR を重視した事業に取り組んでいると考えられる。下位のグループは前期で挙げられた環境問題の対応のため中期に環境技術の活用に取り組んだものの、上位グループ同様に影響を受け、活動の見直しを行っていると考えられる。

輸送の業種では前期以降は完全にコスト比率別で別れた結果となっている。業種の上位に分類された企業が「ディーゼルエンジン」と「トラック」という一貫した目標を持って製品開発を行ってきた結果、下位の結果と大きく距離を開けると同時に比率上位の中期と後期がほぼ同じ位置に存在することとなった。

化学も前期以降にそれぞれのグループがコスト比率の上下で別れた形となった。建設とは逆にコスト比率下位が近く、時系列で整列していない特殊な並びとなっている。また、上位の後期は全体から乖離する動きを見せており、中期付近のキーワードから考察すると開発製造が一段

落し、後期の災害等の影響で活動を大きな転換を図ろうとしていると考えられる。

鉄非鉄も前期以降にコスト比率の上下で別れた形となった。化学に類似した結果であるが、グループ配置が原点に対して円の形に近いため、中期以降のトップメッセージに極端な変化はなかったと考えられる。キーワードの配置をもとにグループの動向を推測すると上位は初期から中期にかけて省エネルギー化やリサイクルを進め、それ以降 CSR への取り組みを深めた。下位は中期から一貫して CSR 方面の取り組みを行っているものの、ほとんど変化が見られない。

家電グループも化学と類似した形でコスト比率の上下で別れた形となったが、下位グループの中期と後期は前期との距離がほぼ同じか後期のほうが前期に近くなっている。キーワードの分布から製品についてコスト比率上位は考え、コスト比率下位は社会責任や CSR を中期に論じ、後期に文化的側面に言及し始めたため期間ごとの並びにならなかったと考えられる。

食品の対応分析では期間別のグループが非常に強く表れた類を見ないグループであり、コスト比率ごとの違いは不明瞭である。逆に言えばこの業種全体が一定の目的を持って活動を行っているといえる。全体の傾向としては中期に責任に関連するキーワードが多く、後期に消費者に関するキーワードが多く配置されている。責任やステークホルダーから消費者と健康へのシフトはより企業の活動がより前向きになったと言える。

制御は前期と中期の壁が他業種と比較して非常に小さいという点が特徴である。そのため、後期が原点から大きく外れるグループ配置となっている。単語の配置から上位グループは製品開発に積極的であり、下位グループはユーザー重視の取り組みを行ってきたと考えられる。

その他業種では、建設同様に明確に前期と中期後期のグループが分かれる珍しい配置となっている。グループ配置から中期以降に下位グループは企業責任の理念から実践への変化、上位グループは CSR への対応から新エネルギーへの取り組みといった活動を行っていたため結果的に下位の後期と上位の中期の活動が結果的に同様のものとなったといえる。

通信では、建設同様に前期と中期後期のグループが分かれており、上位グループが下位グループに合流する形になっている点でも類似している。上位の後期付近に災害を思わせるキーワードが集中していることから、下位の後期もそこにひきつけられていると考えられる。前期は環境活動に取り組み、中期では上位は社会のため環境関連の活動に取り組み、他業種でも見られるように具体性のない下位のグループの活動が、震災をきっかけに大きく変化したと考えられる。

精密でも建設同様に明確に前期と中期後期のグループが分かれており、上位グループが下位グループに合流する形になっている点でも類似している。しかし、キーワードを見ると建設とは逆に上位グループが環境関連のキーワードの近くに配置されており、環境問題に関する関心が高いために中心近くに配置されたと推測できる。また、コスト比率下位の後期のグループ周辺にはキーワードが少なく、具体的活動を示すものもないため、東日本大震災と景気悪化により活動が消極的になった結果と言える。

4.3 まとめ

本研究では、環境報告書に記載されたトップメッセージに対してテキストマイニングを行い、環境保全コスト率の高さと出現頻度するキーワードとの関係を視覚的に検証した。

分析結果から、環境コストを多く支出する企業は、社会とともに CSR や環境問題に対して重大な関心を持ち、新製品の開発や環境対応製品の事業化など積極的な活動を行っており、支出が少ない企業群は、環境関連活動に具体的な目標を持たず、経営改善の一環として CSR 活動を行っている傾向にあることが観察された。

時系列による変化では、CSR の概念の普及に伴って、環境報告書の経営者のメッセージの焦点が地球温暖化対策を中心とした環境問題から環境対応を含む CSR 経営へ大きく変化している。このことは企業による環境活動が社会的責任を果たす活動として広く認知され、環境報告書が企業の広報媒体として活用されるようになってきたといえる。さらに、近年の環境対応製品の事業化により、企業の「環境への対応」の動機と目的は更に複雑化しているといえ、この研究に企業や産業の知識だけではなく、社会や地域といったより広範な知識の重要性が増してきているといえる。

それらの結果から、12年間のトップメッセージのテキスト分析を環境コストの金額という定量的な情報と結び付けてやることで、環境に対する企業の姿勢を直観的な形で浮かび上がらせるという当初の研究目標は達成することができたといえる。

5 . 主な発表論文等

〔雑誌論文〕(計 1 件)

- 1) T. Maeda, Y. Chujo, and E. Park: "Text Mining Analysis on Determinants of Environmental Costs Expenditure as Time Series Data," in Proceedings of The 2018 Technology Innovation Management and Engineering Science International Conference (TIMES-iCON2018), 5 pages (in USB Stick), December 2018. 査読有

〔学会発表〕(計 6 件)

- 1) T. Maeda and M. Yajima: "Motion Frequency Data Analysis for Sports Skill," 2017 Computing Conference (CC 2017), July 2017.
- 2) T. Maeda and M. Yajima: "Proposal of Multi-lingual Sentence Reading for Mobile Communication," International Conference on Computing, Networking and Communications (ICNC 2017), January 2017.
- 3) T. Maeda and M. Yajima: "Sports Skill Discrimination with Motion Frequency Analysis," The 2016 Future Technologies Conference (FTC 2016), December 2016
- 4) T. Maeda, N. Kawakami, Y. Chujo and E. Park: "Time Series Analysis on the Determinants of Environmental Costs Expenditure Using Text Mining Technique," The 2016 International Conference on Parallel and Distributed processing Techniques and Applications (PDPTA 2016), July 2016.
- 5) 米田, 戸谷, 前田: 機械学習を利用した e-ラーニング習熟度評価システム, 情報処理学会第 78 回全国大会, 6ZA-051, 慶應義塾大学(横浜市), 2016 年 3 月
- 6) T. Maeda, Y. Fukushige, and M. Yajima: "Text Mining Analysis for E-Health Information System," 2015 IEEE 17th International Conference on e-Health Networking, Applications and Services (Healthcom 2015), October 2015.

〔図書〕(計 0 件)

〔その他〕

ホームページ等

6 . 研究組織

(1)研究分担者

研究分担者氏名：中條 良美

ローマ字氏名：(CHUJO, Yoshimi)

所属研究機関名：阪南大学

部局名：経営情報学部

職名：教授

研究者番号(8桁): 00387383

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。