

# 有価証券報告書におけるリスク階層構造分析

## Analysis of risk hierarchy structure in securities reports

藤井 元雅<sup>1\*</sup> 坂地 泰紀<sup>2</sup> 佐々木 一<sup>3</sup> 増山 繁<sup>4</sup>

Motomasa Fujii<sup>1</sup> Hiroki Sakaji<sup>2</sup> Hajime Sasaki<sup>3</sup> Shigeru Masuyama<sup>4</sup>

<sup>1</sup> 東京理科大学大学院経営学研究科

<sup>1</sup> Graduate School of Management, Tokyo University of Science

<sup>2</sup> 東京大学大学院工学系研究科

<sup>2</sup> Graduate School of Engineering, The University of Tokyo

<sup>3</sup> 東京大学未来ビジョン研究センター

<sup>3</sup> Institute for Future Initiatives, The University of Tokyo

<sup>4</sup> 東京理科大学経営学部

<sup>4</sup> School of Management, Tokyo University of Science

**Abstract:** In an environment of high uncertainty, risks in corporate management are changing daily. However, risks do not occur independently, but are often caused by a variety of events (risk factors), including other risks. We define this as causal relationships between risk factors. In addition, a single risk factor may have multiple consequences. For example, when a risk factor such as the Covid-19 epidemic occurred in one company, some departments increased sales due to increased IT investment, while others suffered losses as a result of reduced advertising due to worsening business conditions (risk factors of worsening business environment). Therefore, when actually understanding and assessing the risks faced by a company, it is necessary not only to understand the risk factors, but also to understand how they will ultimately affect business performances. In other words, it is important to stratify the risk factors. In this study, we propose a method to obtain the causal content of extracted risk sentences and to understand the risk factors using a hierarchical structure.

## 1. はじめに

昨今、環境・経済・社会がビジネスに与える影響は、多くの企業にとって無視できないリスク要因となっており、SDGs や ESG に貢献する活動を行わないこと自体がリスクになることも指摘されている[1,2]。また、COVID-19 の世界的流行に伴い事業継続計画(BCP)の内容が、企業価値を評価する投資家から重視され始めており[3]、このような不確実性の高い経営環境の中で企業経営におけるリスクは日々変化している。このような事態に対処すべく、藤井ら[4]の研究では深層学習を用いて有価証券報告書に含まれるリスクについて述べた文(以下、リスク文)を抽出・分類する手法を提案した。これは不確実性の高いビジネス環境において、早期のリスク抽出からのリスク管理や企業価値評価に対する重要な視点

を提供している。

しかし、このようなリスクは単独で発生するものではなく、他のリスクを含む様々な事象(リスク要因)が発生したことにより生じる場合が多い。このことをリスク要因間に因果関係があると定義する。さらには1つのリスク要因から複数の別の結果が生じる場合もある。例えばあるIT企業においてはCOVID-19の流行というリスク要因が発生した際に、IT投資の増加によって売上が上がる部門もあれば、景況悪化に伴う広告出稿の抑制(経営環境の悪化というリスク要因)によって赤字に転落する部門もあった。これは、リスク情報を企業価値評価に反映させる上において重要な観点である。例えば「自然災害」をリスクとして抽出したとしても、それがどのように業績に影響を与えるのかを確認しなければ、抽出した情報を企業評価に生かすことができない。あるいは逆に材料費の高騰(コストの上昇)というリスクがなぜ発生するのかを把握しなくては、同様の

\* E-mail: 8621509@ed.tus.ac.jp

問題が生じてくる。つまり、実際に企業が直面しているリスクを把握、評価を行う際には、リスク要因単独を把握するだけでなく、あるリスク要因がまた別のリスク要因を発生させ、それが繰り返されたことで最終的に業績の変動が引き起こされるという一連の流れ(因果関係)を把握する必要がある(図1)。すなわちこれらのリスク要因が波及していく様子を表すためにリスク要因の階層化を図ることが重要であると考えられる。加えて、Campbell et al.[5]はテキストによるリスク開示は、企業が晒されているリスク要因を特定する重要な情報源である点を指摘している。

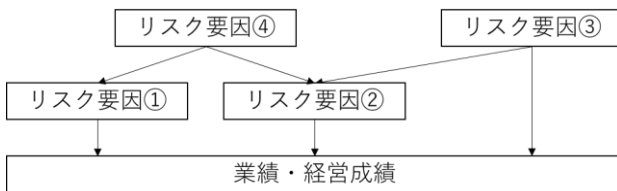


図 1. リスク階層構造の例

本研究では、財務関連書類内のテキストデータからリスク要因の階層関係を取得し、企業価値評価へ役立てることの前段として、実際に人手で作業を行った場合どのような関係が取得できるのか、また取得した情報は企業価値評価において有用であるかどうかを検証する。なお、データソースとしては金融商品取引法で規定された開示資料であり、法令により規格が統一されている有価証券報告書の「事業等のリスク」の項目を利用する。

## 2. 関連研究

テキストデータを対象とする機械学習等によって企業価値評価に貢献する研究は、様々な研究者が行っている。例えば、Huang et al.[6]や Chang et al.[7]は、テキストマイニングを用いた企業価値評価に関する関連研究を行い、その成果として投資判断を行うに際しての企業価値評価の精度向上を確認している。

また、リスクに焦点を当てた研究としては Bao et al.[8]や Dyer et al.[9]があり、いずれも企業のテキストによるリスク開示を分析し、リスク要因を発見する手法を提供している。また藤井ら[4]は、リスクを「行動したことにより獲得できる将来の結果の不確実性」と定義し、該当するリスクが内部要因と外部要因のどちらかによるものかといった複数の観点において、決定木の形で分類することで企業価値評価におけるリスク全てを網羅するリスクの定義と、キーワードなどの既存の手がかりを使わずに財務関連文書等のテキストデータからリスクを抽出及び分類する手法を提案した。手法としては、リスク文でよ

く使われる表現や単語のパターンを手掛かりにするパターンマッチング、SVMなどの教師あり機械学習、BERTなどの深層学習を用いた。その結果、分類及び抽出の両方において深層学習、特にBERTが効果的であることを特定している。この研究においては、リスク要因をそれぞれ並列に独立なものとして抽出及び分類をしているが、本研究においてはリスク間の関係性に焦点を当てている。

また、リスクの階層化の研究として、五十嵐ら[10]は決算短信から因果関係を抽出し、類似性を判断してそれらを接続することでネットワークを構築している。類似性の計算には、日本語 Wikipedia コーパスから作成した word2vec モデルを使用し、単語の重要度を表す idf 値の組み合わせに基づいた手法を用いている。また、因果関係の表現に極性辞書を用いて極性を与え、word2vec が検出できない同義語を併せて判定している。当該研究が一つの事象に対する連鎖的な関係性に焦点を当てているのに対し、本研究においては、企業価値評価に用いるためにリスク要因全体を包括的に網羅する階層関係の作成を試みている。他に Wei et al.[11]は、Sentence Latent Dirichlet Allocation(SentLDA)というテキストマイニング手法を導入し、エネルギー産業のテキストによるリスク開示情報を分析することで、エネルギー企業のリスク要因を包括的に特定するとともに、抽出されたリスク要因の見出しを基に、最終的にエネルギー企業のリスクに影響を与える 66 のリスク要因を特定している。さらに米国のエネルギー企業 840 社を 9 つのエネルギーサブセクターに分けて、リスク要因の階層化を行い、各リスク要因の重要性を出現頻度に基づいて階層化して示している。本研究においては、出現頻度ではなく、リスク要因の因果関係という連鎖の関係性に基いて階層化を試みている。

そして、階層化において重要な要素となる因果関係文の抽出においても様々な研究が行われている。例えば、坂地ら[12]は決算短信から原因・結果表現を自動的に抽出する手法の提案を行っており、具体的には、原因・結果を含む文を抽出した後に、原因・結果表現を抽出した。原因・結果表現抽出には、「～に伴う」「～を背景に」といった手がかり表現を含む構文情報を用いた Pattern を用いている。

本研究においては、有価証券報告書の「事業等のリスク」の項目において、リスク文の中から人手でリスク要因の因果関係を抽出したうえで、リスク階層構造の一部を作成し、その有用性を検討する。

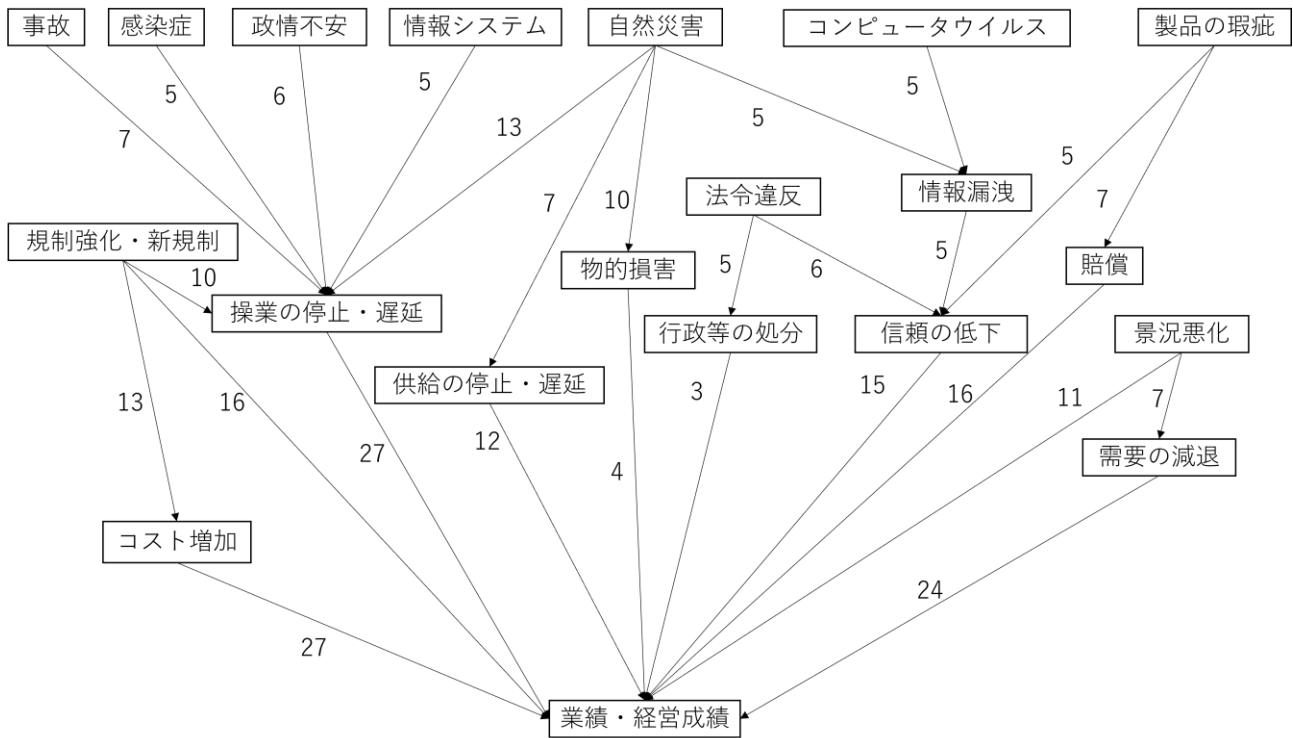


図 2. 抽出したリスク階層構造の一部

### 3. 分析手法

#### 3.1. データ及びデータソース

評価データには、日経 225 に指定されている企業からランダムに 71 社を選択し、それら企業の 2018 年度に発行された有価証券報告書を用いた。有価証券報告書に対しては PDF ファイルをテキスト化し、「事業等のリスク」の項目を抽出したうえで、文単位に分割した。該当する文は計 5,007 文であり人手でタグを付与した結果、リスク文は 2,704 文、リスク文でない文は 2,303 文であった。本研究においては、前者のリスク文と判定された文をデータセットから 500 件を抽出して利用する。なお、リスクの定義については「行動したことにより獲得できる将来の結果の不確実性」としている。

#### 3.2. 人手による因果関係抽出

リスク文への因果関係のタグ付けについては、坂地ら[12]を参考に、以下のように主語(s), 原因(b), 結果(r)の3つに分けて行った。なお、主語については必ずしも文中に含まれているとは限らない。具体例は以下の通りである。

①<s1>当社グループで使用する部材</s1>は<b1>経済環境の変化</b1>により<r1>入手困難になる</r1>可能性があります。  
(日清紡ホールディングス株式会社)

②<b1>市販後に予期していなかった副作用が発生し</b1>た場合には、<r1>当社グループの経営成績及び財政状態等に重大な影響を及ぼす</r1>可能性があります。  
(協和発酵キリン株式会社)

抽出した結果、500 のリスク文から 657 の因果関係を取得した。そして抽出した因果関係に対して人手でクラスタリングを行い、60 のリスク要因に分類をした。以下はその一部である。

感染症 規制強化・新規制 供給の停止・遅延  
競争の激化 景況悪化 コスト増加 法令違反  
コンピュータウイルス 自然災害 信頼の低下  
需要の減退 政情不安 製品の瑕疵 情報漏洩  
操業の停止・遅延 情報漏洩

### 4. 分析結果

前章においてタグ付けしたリスク要因の因果関係は全部で 392 種類確認でき、その中で 3 つ以上確認できた関係を表 1 に示す。また、その因果関係の一

表 1. リスク要因の因果関係例

原因	結果	数	原因	結果	数
コスト増	コスト増	3	事故	操業の停止・遅延	7
コスト増	業績・経営成績	27	事故	物的損害	4
コンピュータウイルス	情報システム	4	自然災害	供給の停止・遅延	7
コンピュータウイルス	情報漏洩	5	自然災害	情報システム	4
シェア低下	業績・経営成績	5	自然災害	情報漏洩	5
マーケティング	業績・経営成績	4	自然災害	操業の停止・遅延	13
為替変動	業績・経営成績	21	自然災害	物的損害	10
価格戦略の失敗	業績・経営成績	8	需要の減退	業績・経営成績	24
株価変動	業績・経営成績	5	需要の減退	原材料高騰	3
株価変動	保有資産の価値低下	4	需要の減退	需要の減退	6
感染症	操業の停止・遅延	5	需要の減退	売上の減少	4
環境汚染	操業の停止・遅延	3	収益性の低下	業績・経営成績	6
機会損失	業績・経営成績	4	行政等の処分	業績・経営成績	3
規制強化・新規制	コスト増	13	情報システム	業績・経営成績	5
規制強化・新規制	規制強化・新規制	3	情報システム	情報漏洩	3
規制強化・新規制	業績・経営成績	16	情報システム	操業の停止・遅延	5
規制強化・新規制	需要の減退	3	情報漏洩	業績・経営成績	8
規制強化・新規制	操業の停止・遅延	10	情報漏洩	信頼の低下	5
技術革新	業績・経営成績	3	信頼の低下	業績・経営成績	15
供給の停止・遅延	供給の停止・遅延	6	人材不足・流出	業績・経営成績	6
供給の停止・遅延	業績・経営成績	12	人材不足・流出	操業の停止・遅延	4
供給の停止・遅延	操業の停止・遅延	4	政情不安	供給の停止・遅延	4
競争激化	価格戦略の失敗	3	政情不安	業績・経営成績	3
競争激化	業績・経営成績	8	政情不安	操業の停止・遅延	6
競争激化	経営経営戦略の不発	5	製品の瑕疵	業績・経営成績	7
競争力低下	業績・経営成績	9	製品の瑕疵	信頼の低下	5
競争力低下	経営経営戦略の不発	3	製品の瑕疵	製品の瑕疵	3
金利変動	業績・経営成績	9	製品の瑕疵	賠償	7
景況悪化	供給の停止・遅延	4	税制	業績・経営成績	3
景況悪化	業績・経営成績	11	経営経営戦略の不発	業績・経営成績	16
景況悪化	需要の減退	7	訴訟	業績・経営成績	14
景況悪化	売上の減少	3	訴訟	賠償	4
研究開発	業績・経営成績	6	操業の停止・遅延	業績・経営成績	27
原材料高騰	コスト増	3	物的損害	業績・経営成績	4
原材料高騰	業績・経営成績	4	貸し倒れ	業績・経営成績	4
原材料高騰	収益性の低下	3	知的財産の侵害	業績・経営成績	7
特定顧客の経営状況	業績・経営成績	3	知的財産の侵害	賠償	4
製品の自主回収・リコール	業績・経営成績	5	投資の失敗	業績・経営成績	4
資金調達の失敗	業績・経営成績	3	売上の減少	業績・経営成績	3
保有資産の価値低下	業績・経営成績	14	賠償	業績・経営成績	16
保有資産の価値低下	保有資産の価値低下	3	価格戦略の失敗	業績・経営成績	3
事故	供給の停止・遅延	4	法令・コンプライアンス違反	行政等の処分	5
事故	情報漏洩	3	法令・コンプライアンス違反	信頼の低下	6
事故	信頼の低下	3	法令・コンプライアンス違反	操業の停止・遅延	3

部を基に作成したリスク階層構造を図2に示す。この図はリスク要因間の因果関係のうち、抽出数が多い関係の一部を抜粋したものであり、矢印の隣の数字は実際に抽出された関係の数である。一部を抽出したこと及び有価証券報告書の記述を基にしており企業が全てのリスク要因を必ず明示するとは限らないことから、ノードへの入力数と出力数の和が一致するとは限らない。例えば、コスト増加には法規制によるもの以外にも、原油価格高騰による燃料コストの増加や人材獲得競争の激化による人件費増加等の様々な要因が考えられる。以下に、これらの階層構造(因果関係)が具体的にどのようなリスク文から抽出されたものかを例示する。

③特に、当社グループの主力商品である 300mm ウェーハの加工工程を有する製造拠点が、上記の自然災害、事故、火災等に見舞われる場合、300mm ウェーハの製造・販売に支障を来し、当社グループの事業展開、経営成績等に影響を及ぼす可能性があります。

(株式会社 SUMCO)

「自然災害」・「事故」  
→ 「操業の停止・遅延」  
→ 「業績・経営成績」

④しかしながら、重大な製品欠陥や製造物責任訴訟の提起といった事象が発生した場合、社会的信用の失墜を招き、顧客に対する補償などによって、当社グループの経営成績及び財務状況に影響を受ける可能性がある。

(昭和電工株式会社)

「製品の瑕疵」  
→ 「信頼の低下」・「賠償」  
→ 「業績・経営成績」

## 5. 考察

これらの内容が最終目的である企業価値評価に有用であるかどうかを検証する。例えば自然災害というリスク要因に焦点を当てた場合、このリスク要因が業績に影響を与えるには、物的損害を与える場合よりも、操業の停止や遅延について言及している(重視している)企業が多いことがわかる。また、自社の操業のみならず、供給ラインも同様に大きなリスク要因であることがわかる。つまり、自然災害発生時には自社への被害や操業への影響だけでなく、供給業者についても同じことが言えるということである。

無論、現在示している内容はある程度推測が容易であり、新たな発見とは言い難い。しかし、リスク要因の因果関係数を測定することで各リスク要因の重要度をマクロ的な視点で推測することができていると言える。また、自然災害において情報漏洩が引き起こされると予測している企業も一部存在している。これはこの一部の企業が他企業に比べ異なる状況にある(ないし、視点を有している)ということであり、このようなミクロの分析においてもこれらの階層構造が新たな視点を提供することができると考えられる。

また、「新たな法規制が当社グループの業績に影響を与える可能性がある」といったリスク文と図に示した構造を照らし合わせることで、コスト増加というリスク要因が有価証券報告書に記載されていなかったとしても、そのリスク要因が当該企業に存在する可能性があることが推測できる。このように、リスク階層構造は抽出されたリスク要因を企業価値評価に生かすに際して、有用であると考えられる。

## 6. まとめ

本研究においては、財務関連書類内のテキストデータからリスク要因の階層関係を取得し、企業価値評価へ役立てることの前段として、実際に人手で作業を行った場合どのような関係が取得できるのか、また取得した情報は企業価値評価において有用であるかどうかを検証した。結果として、抽出されたリスク要因の因果関係によって構成されたリスク階層構造の情報は企業価値評価に有用であるということを確認することができた。しかし今回は人手による作業であるためデータセットが少なくなっているが、実際に実用に足るデータを用意するには大量のデータを処理する必要があり、そのためには機械学習や深層学習が必須であると言える。

上記の研究を進めるにあたって考慮すべきはリスク文抽出、因果関係抽出の2点である。まず、リスク文抽出作業は財務関連書類のテキストデータから不要な文を削除するために必要である。これは藤井ら[4]が示したように人手でタグ付けをした後に双方向 LSTM[13]あるいは BERT[14]を用いて文の分類作業を行えば問題ないと考えられる。次に、因果関係抽出については坂地ら[12]が示した手がかり表現や構文パターンを基にリスク文に焦点を当てた手法に調整をする必要があると考えられる。

## 参考文献

- [1] 経済産業省: SDGs 経営ガイド, 2019.
- [2] Kim, MinChung, and YongHee Kim. "Corporate social responsibility and shareholder value of restaurant firms." *International Journal of Hospitality Management* 40 (2014): 120-129.
- [3] Mukherjee, Mahua, et al. "Ecosystem-centric business continuity planning (eco-centric BCP): A post COVID19 new normal." *Progress in Disaster Science* 7 (2020): 100117.
- [4] 藤井元雅, 坂地泰紀, 佐々木一, 増山繁, 有働証券報告書からのリスク文抽出の試み, 25 回人工知能学会 金融情報学研究会, pp.44-48, 2020.
- [5] Campbell, John L., et al. "The information content of mandatory risk factor disclosures in corporate filings." *Review of Accounting Studies* 19.1 (2014): 396-455.
- [6] Huang, Win-Bin, et al. "Value assessment of companies by using an enterprise value assessment system based on their public transfer specification." *Information Processing & Management* 57.5 (2020): 102254.
- [7] Chang, Jun, et al. "Assessing dynamic qualities of investor sentiments for stock recommendation." *Information Processing & Management* 58.2 (2021): 102452.
- [8] Bao, Yang, and Anindya Datta. "Simultaneously discovering and quantifying risk types from textual risk disclosures." *Management Science* 60.6 (2014): 1371-1391.
- [9] Dyer, Travis, Mark Lang, and Lorien Stice-Lawrence. "The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation." *Journal of Accounting and Economics* 64.2-3 (2017): 221-245.
- [10] 五十嵐光秋, et al. "極性を考慮したリスク発見に向けた因果関係ネットワークの構築." 人工知能学会全国大会論文集 第 33 回全国大会 (2019). 一般社団法人 人工知能学会, 2019.
- [11] Wei, Lu, et al. "Developing a hierarchical system for energy corporate risk factors based on textual risk disclosures." *Energy Economics* 80 (2019): 452-460.
- [12] 坂地泰紀, 酒井浩之, 増山繁, 決算短信 PDF からの原因・結果表現の抽出, 電子情報通信学会論文誌 D, Vol.J98-D, No.5, pp.811-822, 2015.
- [13] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." *Neural computation* 9.8 (1997): 1735-1780.
- [14] Devlin, Jacob, et al. "Bert: Pre-training of deep bidirectional transformers for language understanding." *arXiv preprint arXiv:1810.04805* (2018).