

AI 支援によるサステナビリティ指標の試行的取り組み －SDGs 時代における社会的評価手法の開発への挑戦－

中尾悠利子（公立鳥取環境大学），石野亜耶（広島経済大学），岡田齋（広島経済大学），國部克彦（神戸大学）

1. はじめに

本研究では、サステナビリティ情報（企業の環境・社会・ガバナンス（ESG）情報を『サステナビリティ情報』と呼ぶ）の記述情報（自然言語の記述情報を以下『テキスト』と呼ぶ）に対して AI による自然言語処理技術を用いることにより、ビッグデータへの適用を可能とした新しい社会的評価手法の開発に挑戦する。

本研究の意義は、まず、AI による独自の自然言語処理をサステナビリティ情報に適応することにより、文脈から切り離された単語単位の分析ではなく、文脈を含めた分析を行う点であり、次に、AI 支援による自然言語処理技術により、ビッグデータへの適応が可能となる点、さらに、これらの処理技術の開発により、財務情報を中心とした市場において、サステナビリティ情報を含めた企業の社会的評価手法の開発が可能になる点があげられる。

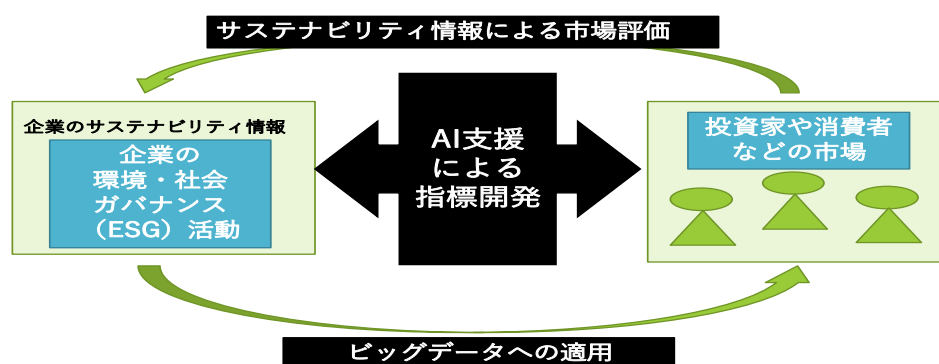


図1 AI 支援による指標開発とサステナビリティ情報の市場評価との関係

わが国では企業の環境・社会活動は自主的に行われている。したがって、そのような環境・社会活動は、企業の自由裁量により行われるため、経営トップによる環境・社会活動のリーダーシップは重要との指摘がある(e.g., Waldman, 2006)。それでは、わが国の経営トップは環境・社会活動は積極的なのだろうか。Harvard Business Review (HBR) (2017) の調査によると、わが国経営トップの環境・社会活動の取り組み指標は、他国の CEO よりも低い値が示されている。つまり、わが国経営トップの環境・社会活動への取り組み姿勢は他国よりも低いことが考えられる。したがって、わが国の経営トップによる環境・社会活動の取り組み姿勢を検討することは重要な研究課題といえる。そこで本報告では、サステナビリティ報告の定性情報の中でも、最も重要な経営トップメッセージの記述全体に占める環境・社会

の記述割合の分析を AI 支援により行い、経営トップの環境・社会活動への意図を検討する。

2. 実験方法の概要

本研究では、サステナビリティ報告の 1 文に対し、その文が環境について述べている文か、社会について述べている文かを、機械学習により自動判定する手法を提案する。ここでは、環境について述べている文には環境ラベル、社会について述べている文には社会ラベルを付与する問題とする。機械学習には、ニューラルネットワークの一種である Deep Averaging Networks (DAN) (Iyyer et al. (2015)) を利用する。

2. 1 実験に使用したデータ

実験には、サステナビリティ報告に対し 1 文ごとに人手で正解ラベルの判定を行った結果を利用した。正解ラベルの判定は 2 名の専門家¹により行った。正解ラベルを判定するデータは、2,000 以上の文章を確保するため、サステナビリティ報告 70 件、2,188 文を対象とした。人手で正解ラベルを判定した結果を表 1 に示す。

表 1 サステナビリティ報告を専門家で分類した結果

環境ラベル	社会ラベル
987	984

2. 2 既存の分析ツールとの比較

本研究で開発した機械学習モデルによる環境・社会分類と既存のテキストマイニングソフト²から抽出された単語による環境・社会分類³との比較を行う。

2. 3 評価尺度

5 分割交差検定により評価実験を行った。また、評価尺度として、以下に示す精度、再現率、F 値を用いた。

- 精度 Precision = $\frac{\text{機械が正しくラベルを付与した件数}}{\text{機械がラベルを付与した件数}}$
- 再現率 Recall = $\frac{\text{機械が正しくラベルを付与した件数}}{\text{人手でラベルを付与した件数}}$

¹ 専門家は本研究の従事者中尾および岡田により実施した。中尾は、民間会社にてサステナビリティ報告のコンサルティング業務に 10 年近く従事した経験を有する。岡田は、サステナビリティ報告の第三者意見など、専門的立場からの助言を多数行っている。

² 本研究では、学術領域で多く利用されている IBM SPSS Text Analytics for Surveys 4.0 のソフトを採用した。

³ テキストマイニングソフトから抽出された、環境に分類した単語は、533。他方、社会は、115。環境および社会の両方に該当する単語は 115 となった。

● F 値
$$F - \text{measure} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

2. 4 実験結果

既存の分析ソフトの単語をもとにした「環境」分類および「社会」分類と、本研究で開発した機械学習モデルによる「環境」分類および「社会」分類の実験結果は表 2 の通りである。「環境」の精度「0.819」は、既存の分析ソフトが高い値を示している。しかし、その他の「環境」および「社会」すべての評価指標は、本研究で開発した機械学習モデルの方が高い値を示している。

表 2 実験結果

手法	環境			社会		
	精度	再現率	F 値	精度	再現率	F 値
既存分析ソフトの単語分類	0.819	0.403	0.54	0.699	0.403	0.513
開発した機械学習モデル	0.723	<u>0.702</u>	<u>0.713</u>	<u>0.723</u>	<u>0.691</u>	<u>0.706</u>

3. モデルの適用：経営トップメッセージの環境・社会割合（2001年～2017年）

本研究では 1 文の文脈をもとにサステナビリティ報告の文章を「環境」、「社会」の分類分けする機械学習モデルを開発した。次に、本研究ではその開発したモデルを、サステナビリティ報告の 2001 年から 2017 年を対象に「918」の経営トップメッセージ、「29,669 文」に適用した⁴。2001 年から 2017 年全体の環境・社会の文章の数および経営トップメッセージ全体の文章量に対する環境および社会割合の記述統計量は、次の表 3 の通りである。図 2 は、2001 年から 2017 年までの環境および社会割合の各年の平均値(54 社)を折れ線グラフで示している。図 3 は、図 2 と同様に、2001 年から 2017 年までの環境および社会文章割合の各年の平均値(54 社)を折れ線グラフで示している。

表 3 環境・社会文章の数および、環境・社会割合(2001年～2017年)

記述統計量	平均	最大	最小	標準偏差	N (サンプルサイズ)
2001年～2017年					
文の数	32.32	158	1	21.82	918
環境の文の数	12.89	67	0	8.04	918
社会の文の数	13.31	67	0	9.73	918
環境の割合	0.44	1.00	0.00	0.19	918
社会の割合	0.41	0.91	0.00	0.19	918

⁴ 2015 年の売上高上位 100 社の中で、2001 年から 17 年間継続してサステナビリティ報告を発行している企業 54 社、918 社を対象とした。

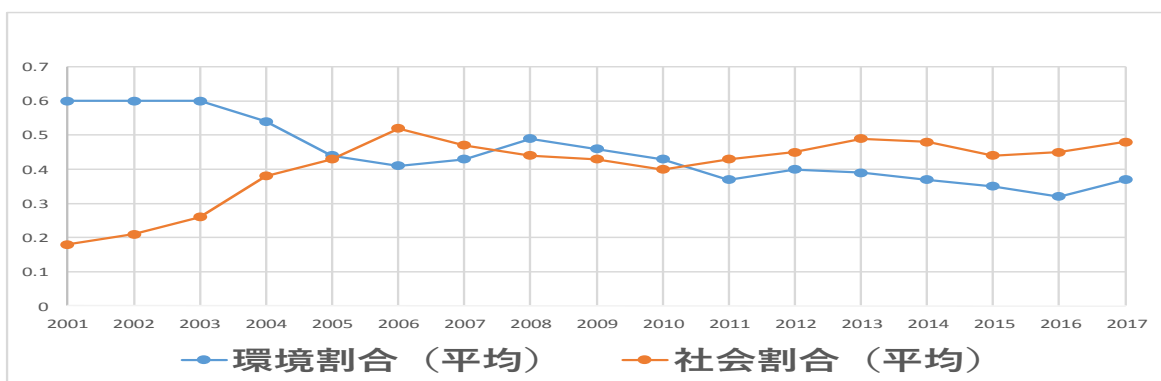


図2 環境および社会割合の各年の平均値(54社) : 2001年～2017年

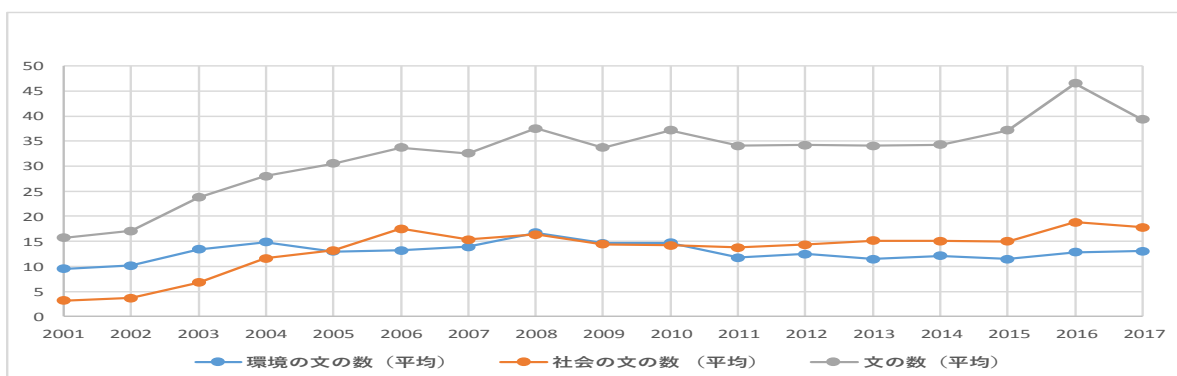


図3 環境および社会文章の各年の平均値(54社) : 2001年～2017年

表4の結果、および図2、図3の結果は一部ではあるがこれらの結果は、機械学習モデルをサステナビリティ報告のテキストに適用することで、多くの文章を対象とした企業の環境・社会開示傾向が可能になることを示した。本報告で開発したモデルにより、従来のサステナビリティ情報の実証研究の限界点（人手による分析のため限られた観測数）を克服し、企業の裁量的情報開示行動を研究する上での新しい展開が期待される。

参考文献

- Harvard Business Review Staff (2017) , The Best-Performing CEOs in the World, https://hbr.org/2017/11/the-best-performing-ceos-in-the-world-2017?referral=03758&cm_vc=rr_item_page.top_right (access to 19, October, 2018)
- Mohit Iyyer, Varun Manjunatha, Jordan Boyd-Graber, Hal Daumé III, Deep Unordered Composition Rivals Syntactic Methods for Text Classification, Proceedings of ACL2015.
- Waldman, D. A., Siegel, D. S., & Javidan, M. (2006). Components of CEO Transformational Leadership and Corporate Social Responsibility. *Journal of Management Studies*, 43(8), 1703-1725.