

はじめに～重要性を増す ESG 関連情報をどう分析するか

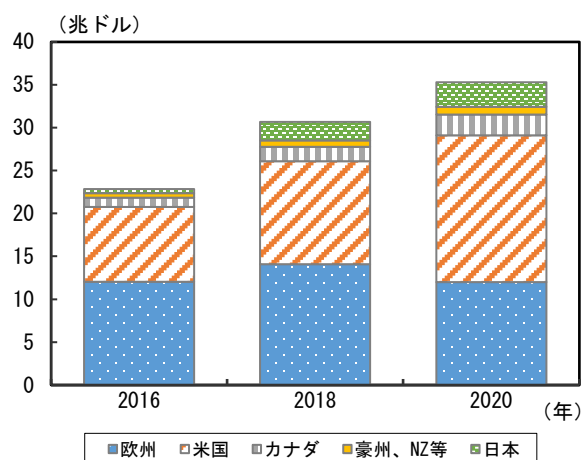
近年、ESG 投資の急速な拡大を受けて、企業が公表する ESG 関連情報への関心が高まっている。世界持続可能投資連合（GSIA、Global Sustainable Investment Alliance）の集計によると、世界主要5ヵ国・地域（日本、米国、欧州、カナダ、豪州・NZ等）のESG投資残高は、2020年時点で約35.3兆ドルとなっている（図表1）。特に、日本における投資残高は2016年から2020年にかけて6倍に急増した。

こうした中、これまで主に任意開示の形で公表されてきた ESG 関連情報に対して、その開示を制度化・義務化する動きが進んでいる。国内では、有価証券報告書やコーポレート・ガバナンス報告書において ESG 関連情報の開示が要求されており、今後に向けた開示内容の拡充も議論の最中にある。企業側も積極的に開示する動きを見せており、企業の持続的な価値創造に関する重要な財務情報と ESG を含む非財務情報を一つの媒体で伝える「統合報告書」を公表する企業が増加している。企業価値レポート・ラボ（株式会社エッジ・インターナショナルが運営）の調査によると、2021年に統合報告書を公表した企業等は716社となり、前年（599社）から117社も増加した（図表2）。

企業の ESG 関連情報開示が拡大していく中で、その情報をどのように分析するかが新たな課題の一つとなっている。非財務情報である ESG 関連情報は、数値データではなく文字データとして公表されている部分も多い。また、統合報告書は、一般的に1社あたり数十ページに及ぶ文書となっており、業界の全体的な動向を掴むために数十社分の資料を読むことになると、多大な時間が必要となる。

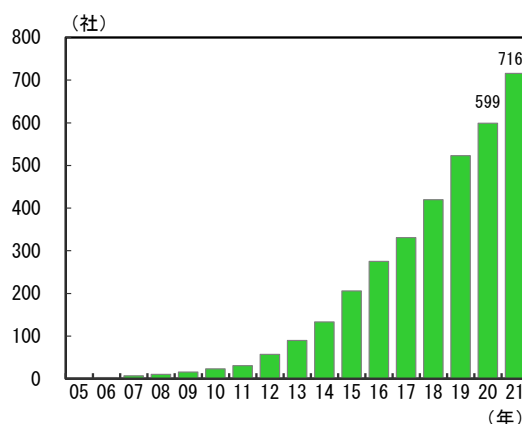
そこで、本稿では、テキストマイニング及び機械学習を活用し、効率的かつ視覚的に特定業界の統合報告書の全体的な動向を掴むことを試みる。今回は銀行業を対象にして、その統合報告書の公表状況を概観するとともに、銀行種別に ESG 関連のどのようなトピックを重視しているかを可視化し、今後の注目点や課題について検討する。

図表1 世界の ESG 投資残高推移



(注) 日本以外は表示年前年末現在。
日本は表示年3月末現在。
(出所) GSIA より大和総研作成

図表2 統合レポート発行企業等の推移



(注) 統合レポートは、自己表明型統合レポートで企業価値レポート・ラボ調べ（発行年ベース、2021年は2022年1月20日発行分までをカウント）。
(出所) 企業価値レポート・ラボ「国内自己表明型統合レポート発行企業等リスト 2021年版」より大和総研作成

1. 銀行業における統合報告書の公表状況とその特徴

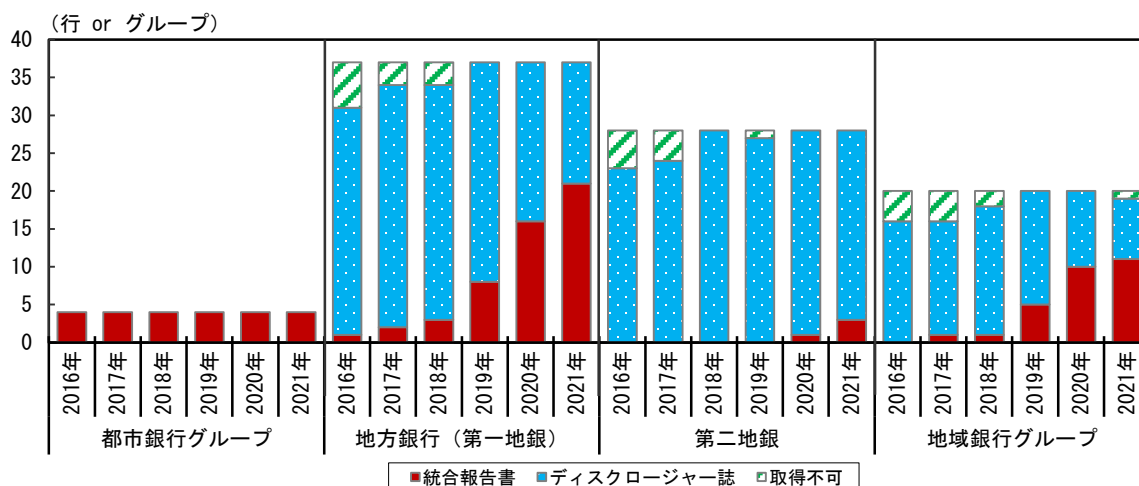
(1) 銀行種別のディスクロージャー資料の開示形態

銀行業における統合報告書の公表状況と銀行種別の特徴等を整理する。図表3は、ウェブサイトに掲載されている2016年（各年は公表日ベース）以降の銀行業のディスクロージャー資料¹について、その開示形態を銀行種別に時系列で示したものだ（大和総研調べ）。

都市銀行グループにおいては、全てのグループがこの間一貫してディスクロージャー資料を統合報告書として公表している。地方銀行（以下、第一地銀）²は、2016年時点で統合報告書による公表割合は非常に低かったものの、近年明確な増加傾向にあり、2021年時点で過半となっている。第二地銀においては、統合報告書の公表が進んでおらず、依然として、従来型のディスクロージャー誌が中心となっている。地域銀行グループ（第一地銀や第二地銀を保有する銀行グループ）は、第一地銀と同じように、近年、統合報告書の公表を進めてきたことが分かる。

このように銀行業の統合報告書の公表状況は、全体として増加傾向にあるものの、銀行種別に見ると、大きく異なっている。いまだ統合報告書を公表していない銀行（含む銀行グループ、以下同様）にとっては、その作成負担が小さくないことや、短期的なメリットを実感しにくいことなどが、二の足を踏む要因になっていると考えられる。ただし、先行きについては、株主をはじめとする様々なステークホルダーとの対話強化や自行の価値創造プロセスの明確化といった長期的なメリットなどを踏まえ、第一地銀と地域銀行グループにおいては今後も公表拡大の動きが着実に進み、第二地銀による公表も徐々に増加すると見込まれる。

図表3 銀行業におけるディスクロージャー資料の開示形態の推移（ウェブ掲載ベース）



(注1) 統合報告書はウェブサイトの開示資料のうち、その旨明記されているもの（大和総研調べ）。

(注2) 取得不可は統廃合などにより継続的な取得が不可能なものや、取得時点でウェブサイトに掲載されていないもの。

(注3) 地域銀行グループ内の銀行は第一地銀や第二地銀の項目に計上していない。

(出所) 各行のウェブサイトの開示資料より大和総研作成

¹ ディスクロージャー資料は銀行法により作成・公表が義務付けられている。

² 本稿においては、全国地方銀行協会に所属する地方銀行を「第一地銀」、第二地方銀行協会に所属する地方銀行を「第二地銀」と呼称する。

(2) 上場銀行の属性と統合報告書の公表割合

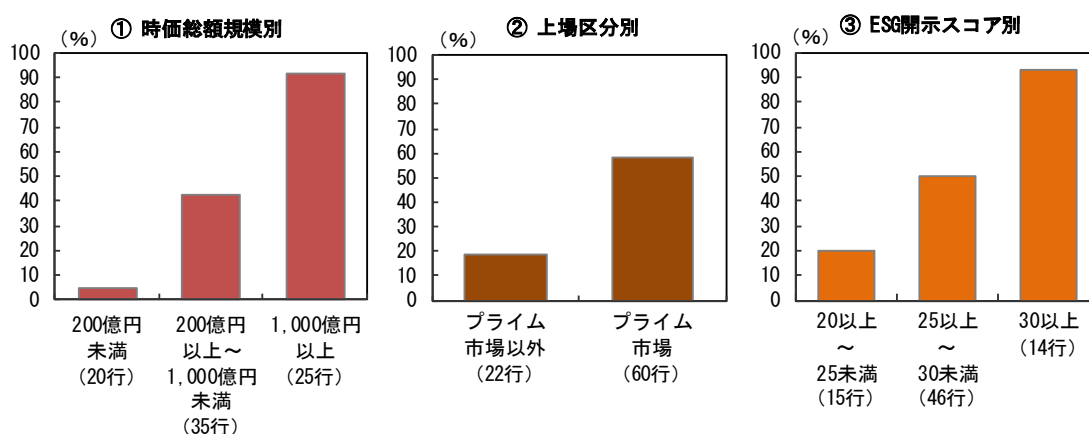
統合報告書の公表実態を多面的に捉えるために、銀行業のうち上場銀行（含む銀行グループ、以下同様）を対象として、①時価総額規模別、②上場区分別、③ESG 開示スコア別、に見た 2021 年の統合報告書の公表割合を示したのが図表 4 だ。

第一に、時価総額別に見ると、時価総額（2021 年度末）が 200 億円未満の銀行において、統合報告書の公表がほとんど進んでいない。なお、この区分には第二地銀が多く含まれている。他方、時価総額が 1,000 億円以上の銀行では、9 割以上が統合報告書を公表している。この結果から、規模の小さな銀行において、統合報告書の作成に関わる負担等が相対的に重くなっており、それが公表の妨げになっている可能性が示唆される。

第二に、東京証券取引所（東証）の上場区分別に見ると、最上位のプライム市場に上場している銀行はすでに 5 割以上が統合報告書を公表している一方、それ以外の銀行は 2 割程度の公表にとどまっている。プライム市場の上場企業は、2021 年 6 月に改訂された「コーポレートガバナンス・コード」で気候関連財務情報開示タスクフォース（TCFD、Task Force on Climate-Related Financial Disclosures）に基づく事業リスク等の開示が要求される³など、グローバルな投資家との建設的な対話を念頭に、より高い水準での情報開示が期待されている。このため、最上位のプライム市場に上場するような銀行ほど積極的に統合報告書を公表していたとみられる。

第三に、Bloomberg の「ESG Disclosure Score」（ESG 開示スコア、2021 年度）との関係を確認すると、ESG 開示スコアが高い（低い）ほど統合報告書の公表割合が高い（低い）という正の相関関係が見られ、ESG 開示スコアが 30 以上の場合、9 割以上が統合報告書を公表している。なお、ここで注目したいのは、統合報告書を公表しているにもかかわらず ESG 開示スコアが低い銀行だ。そうした銀行は、統合報告書の記載事項等に改善の余地があると推測される。

図表 4 上場銀行の属性別に見た統合報告書（2021 年）の公表割合



(注 1) 上場銀行（含む銀行グループ）のうちデータが取得できたところのみを集計。括弧内が集計行数。

(注 2) 時価総額は 2021 年度末、ESG 開示スコアは Bloomberg の「ESG Disclosure Score」で 2021 年度。

(出所) 各行のウェブサイトの開示資料、日本取引所グループ、Bloomberg より大和総研作成

³ 具体的には、「(前略) 国際的に確立された開示の枠組みである TCFD またはそれと同等の枠組みに基づく開示の質と量の充実を進めるべきである。」(コーポレートガバナンス・コード 補充原則 3-1③) とされた。

2. テキストマイニングによる統合報告書等のホットトピック可視化

(1) 分析の対象

本稿のテキストマイニング及び機械学習を活用した分析では、ウェブサイトから「統合報告書」または「ディスクロージャー誌」を PDF ファイルの形式で取得できる銀行を対象とした。実際に分析対象とした銀行数は、2016 年が 74 行/グループ、2021 年が 87 行/グループである。

(2) 分析の流れ

今回の分析は、①データ収集、②データ整備、③ワードクラウド⁴による可視化、の3段階で行った(図表5)。各段階の詳細については、次の通りである。

①データ収集

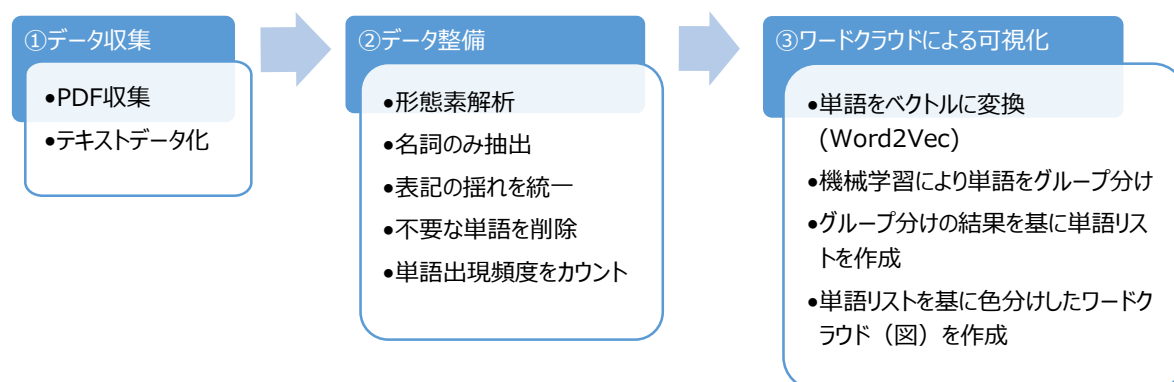
各行のウェブサイトから統合報告書またはディスクロージャー誌の PDF ファイルを収集して、テキストデータに変換した。

②データ整備

テキストデータを最小単位の単語ごとに分割(形態素解析)した。分割した単語群から名詞(以下、単語)のみを取り出し、表記に揺れがある場合は表記を統一(例:サステナビリティ、サステナビリティー、Sustainability は、全てサステナビリティに統一)した。その後、分析に不要と思われる単語(ストップワード)を削除した。不要な単語とは、多用される一般的な単語(例:年度、月)や財務諸表中に記載される会計項目名等である。

最後に、次の段階でワードクラウドを作成するため、残った単語群の出現頻度をカウントした。なお、カウントする際には、極端なカウントデータに分析結果全体が左右されることを防ぐため、銀行ごとに同一単語のカウント上限を20に設定し、かつ一つの銀行の開示資料でしか使われていない単語は集計の対象外とした。

図表5 分析の流れ



(出所) 大和総研作成

⁴ ワードクラウド (Word Cloud) とは、文書中出现した単語を集めて、その出現頻度に応じて文字の大きさを変えたり色分けしたりして、文書全体の特徴を見やすく可視化する手法や、可視化した図のことである。出現頻度が高い単語ほど大きな文字で表示される。

③ワードクラウドによる可視化

ここでは、統合報告書の文書の特徴を分かりやすく可視化するために、単語（名詞）をその内容に応じてグループ化した上で、各グループを色分けして、ワードクラウド（図）を作成した。技術的な詳細については省略するが、概要は次の通りである。

まず、「Word2Vec」という手法により単語をベクトル（大きさと向きを持つ量）に変換した。次にクラスタリングという機械学習の手法によりグループ分けを行った。グループ分けした結果を基に人の手による調整を行った上で、単語リストを作成した。最後に、単語リストをグループごとに色分けしてワードクラウド（図）を作成した。

具体的な色分けとして、「環境関連（E）」：緑色、「社会関連（S）」：桃色、「ガバナンス関連（G）」：紫色、「ビジネス関連」：青色、「金融関連」：黄色、「新型コロナウイルス関連」：茶色、「デジタル関連」：赤色、「役職関連」：水色、「その他」：灰色としている。なお、SDGs やサステナビリティ等は ESG 全体と広く関わる単語であるが、分類の必要性から今回は「環境関連」とした。

（3）分析結果～都市銀行は環境関連、第一地銀は社会関連への関心が強い傾向

今回のワードクラウドによる分析結果は、次頁図表 6 で示した。また、その見方を簡単に述べると次の通りである。

まず、図表の上段に都市銀行グループ、中段に第一地銀、下段に第二地銀を示した。なお、地域銀行グループのディスクロージャー誌または統合報告書（以下、統合報告書等）は、傘下に複数の銀行が存在しても一つとなる。そこで、傘下の銀行が全て第二地銀で構成される場合は第二地銀として扱い、それ以外の場合は第一地銀とした。

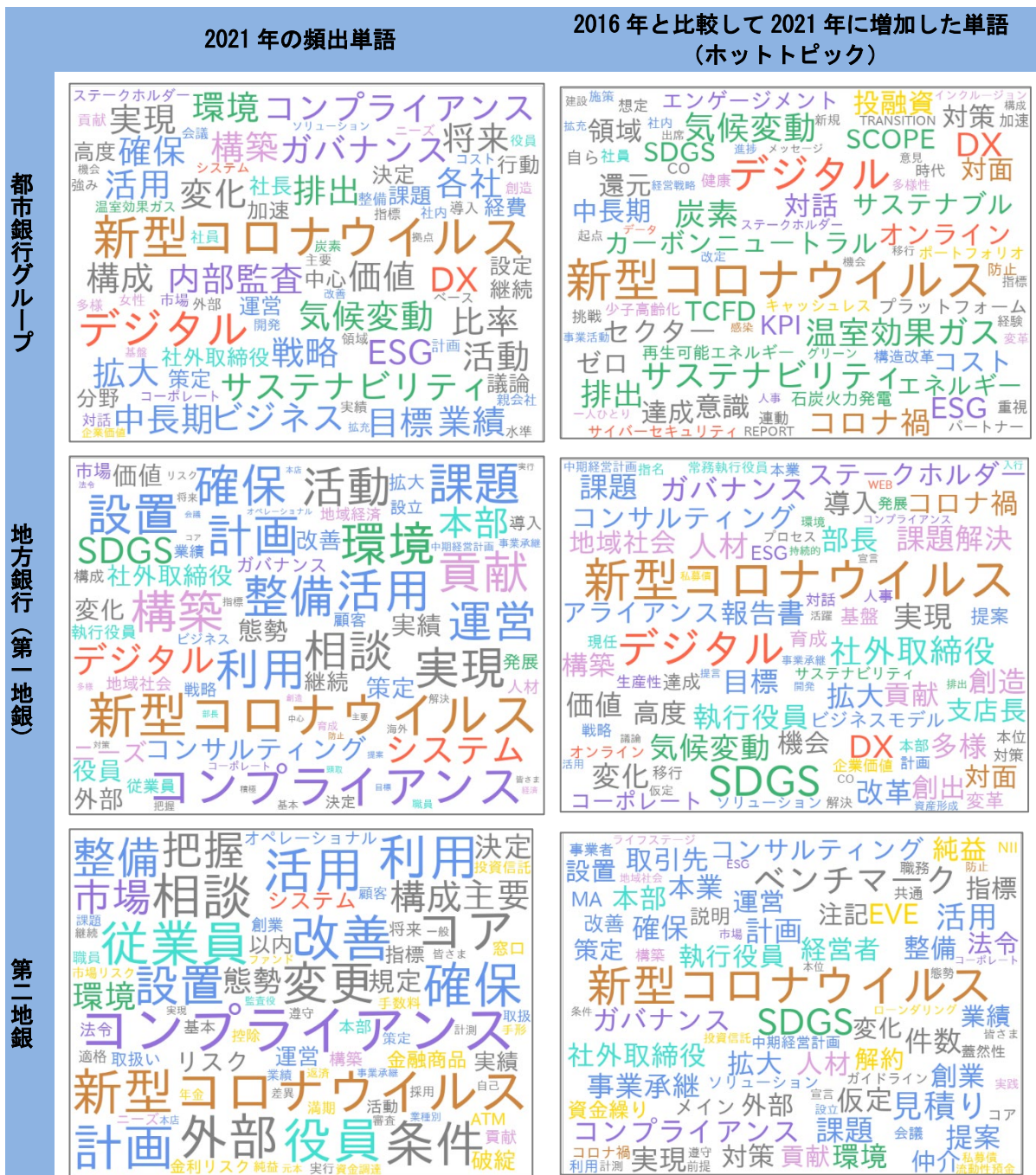
次に、左側の列は、2021 年の統合報告書等を基に作成しており、その中で出現頻度の高い単語ほど大きく表示している。右側の列は、各社の 5 年前（2016 年）の統合報告書等と比較して増加した単語ほど大きく表示するように図示したものであり、直近（2021 年）の開示資料における「ホットトピック」（注目の話題）とした。

全ての銀行種に共通する頻出単語としては、未曾有の厄災という外生要因の影響により、「新型コロナウイルス」が最も目立っており、銀行業全体への影響の大きさがうかがえる。銀行種別の特徴については、以下で確認していく。

① 都市銀行グループ～環境とデジタル対応を一層重視

2021 年の頻出単語を見ると、多岐にわたる ESG 関連の単語（緑色、桃色、紫色）が図示されていることが大きな特徴であり、企業の情報開示項目や経営課題として ESG 分野を幅広く重視している様子が読み取れる。他には、金融 DX（デジタル・トランスフォーメーション）の潮流を背景に、デジタル関連（赤色）の単語の大きさが目立ち、さらにビジネス関係（青色）の単語も多く登場している。

図表6 銀行の開示資料（ディスクロージャー誌または統合報告書）における頻出単語



(注1) ディスクロージャー誌または統合報告書をPDFファイルで取得できる銀行を対象とした。地域銀行グループの開示資料は、傘下銀行が第二地銀のみで構成されるものを第二地銀に分類し、その他の場合は地方銀行（第一地銀）に分類している。対象銀行/グループ数は2016年が74行/グループ（サンプル数内訳：都市銀行4、第一地銀46、第二地銀24）、2021年が87行/グループ（サンプル数内訳：都市銀行4、第一地銀53、第二地銀30）である。

(注2) 環境関連：緑色、社会関連：桃色、ガバナンス関連：紫色、ビジネス関連：青色、金融関連：黄色、新型コロナウイルス関連：茶色、デジタル関連：赤色、役職関連：水色、その他：灰色としている。なお、SDGsやサステナビリティ等はESG全体と広く関わる単語であるが、分類の必要性から今回は環境関連とした。

(出所) 各行のウェブサイトの開示資料より大和総研作成

2016年との比較では、特に環境関連の単語（緑色）が幅広く増加していることが注目される。この背景の一つには、TCFDが2017年6月に最終報告書（TCFD提言）を公表して以降、国内外で、その提言に沿った気候変動関連の情報開示が進められてきたことがある。環境関連の単語について銀行種別の結果を比較すると、第一地銀と第二地銀より、都市銀行グループにおいて気候変動関連の情報開示が先行してきたことが示唆される。また、東証が、プライム市場の上場企業に対してTCFDに基づいた気候変動関連の情報開示を要求しており、「気候変動」や「TCFD」といった単語が2022年以降、さらに増加するかが今後注目されよう。

他には、デジタル関連の「DX」や「デジタル」という単語が増加しており、業務の効率化や収益性の改善などのために、金融DXの取り組みを積極的に推進している様子が分かる。また、「プラットフォーム」という単語が増加している点に注目したい。近年、都市銀行グループは、金融DXを起点に、金融プラットフォーム・ビジネスを強化して収益機会の拡大を図っており、こうした経営戦略における重点テーマもしっかり捉えられている。

②第一地銀～地域社会・経済への貢献という意識の高まり

2021年の主な特徴としては、都市銀行グループと同様に、ESG関連（緑色、桃色、紫色）やデジタル関連（赤色）、ビジネス関連（青色）の単語において、文字サイズの大きなものが見られ、単語の種類も多い。他方、ESG関連のうち社会関連（桃色）において、文字サイズはやや小さいものの、「地域社会」や「地域経済」といった単語が登場している点が都市銀行グループと異なる。地元の地域社会・経済により密着した経営形態を持つ地方銀行の特徴が表れているといえるだろう。

2016年との比較を見ると、環境関連（緑色）において、「SDGs」や「気候変動」といった大枠のテーマの単語が増加している一方、環境関連の単語の種類は都市銀行と比べてかなり少ない。このことは、第一地銀全体として見ると、環境問題の各論にまで踏み込んだ議論がまだ十分行われていない可能性を示唆する。

なお、今回の分析対象となった都市銀行グループの開示資料は全て統合報告書であるのに対し、第一地銀は、2021年にかけて従来型のディスクロージャー誌から統合報告書に切り替えるところが増えてきた（前掲図表3）。そのため、事前の想定では、統合報告書への切り替えの影響により、第一地銀において環境関連の単語が目立って増えるとみていたが、結果として都市銀行グループよりも限定的なものとなった。今後については、東証のプライム市場に上場している第一地銀を中心に、「気候変動」や「TCFD」をはじめとする環境関連の単語が増えていくか注目していきたい。

他には、都市銀行グループと比較して、第一地銀ではESG関連のうち「地方社会」や「貢献」など社会関連の単語（桃色）が増えていることが目立つ。地方創生の実現に向け、地域金融の分野から地域社会・経済に貢献していくという意識が高まっている様子がうかがえる。

③第二地銀～デジタル関連の単語の少なさが気になり

2021年について確認すると、ESG関連の単語（緑色、桃色、紫色）よりもビジネス関連の単語（青色）が多いことが目立つ。この背景の一つとしては、今回の分析対象となった第二地銀の開示資料のほとんどが統合報告書ではなくディスクロージャー誌であったことが挙げられる。銀行によっては、ESG関連の取り組みをCSRレポートに別途記載して、ディスクロージャー誌にあまり記載していないケースもあろう。また、デジタル関連については、「システム」といった単語が見られる程度であり、都市銀行グループや第一地銀ほど強調されていないようである。

2016年との比較で見ると、環境関連の単語（緑色）は増加傾向にあるが、第一地銀と同様に、「SDGS」や「環境」といった大枠のテーマの単語であり、開示資料で各論にまで踏み込んだ言及はなされていないようである。また、全体的に増加している単語は、2021年の結果からもうかがえるように、事業関連の単語が多い。

他方、都市銀行グループや第一地銀と異なり、デジタル関連の単語は増えていない。現在、銀行業では、業務の効率化や収益性の改善のために金融DXの取り組みが重要な課題となっているが、第二地銀の開示資料からはそうした様子が見えないう結果となった。第二地銀における金融DXの動向は、株主等のステークホルダーにとって関心の高いテーマの一つであると考えられることから、第二地銀には、今後の開示資料でデジタル関連の情報を拡充していくことなどが求められよう。

（４）【参考】可視化向上の工夫について

ワードクラウドによる統合報告書等の可視化では、クラスタリング（グループ分けを行う機械学習の手法）の結果を参考に単語をグループ化した上で色分けし、文章の特徴を視覚的により分かりやすく示すという工夫を行った。実際のグループ化においては、人の手による調整が必要なケースもあった。例えば、ESGの環境関連の単語とガバナンス関連の単語が一つのグループに混在するケースが見られた。他方、金融関連の単語や役職関連の単語など、具体的に使用される分野が明確な単語であるほど、筆者の感覚としても適切にグループ化を行うことができおり、ワードクラウドの視認性向上という点において一定程度有効であったといえよう。

おわりに

ESG投資の拡大という大きな潮流の下で、今後も統合報告書を公表する企業は増加していくと見込まれる。こうした中、統合報告書の読み手側となる株主等のステークホルダーと書き手側の企業担当者の双方にとって、個別企業と業界全体の統合報告書の特徴や直近のホットトピックなどを適切に把握することが重要な課題となってくるだろう。

その一つの新しいアプローチとして、本稿では銀行業を対象にしてワードクラウドによる可視化を行い、銀行種別に見たESG分野等に対する意識の違いやホットトピックの抽出を試みた。具体的な分析結果として、都市銀行グループでは、環境問題への意識の強さやデジタル化対応

を重視する姿勢をうかがうことができた。第一地銀においては、地域社会・経済への貢献を重視する姿勢が読み取れた。第二地銀は、従来型のディスクロージャー誌による開示がほとんどであることなどからビジネス関連の単語が多くなっている一方、他の銀行種別に比べてデジタル関連の単語が少ないという特徴が見られた。

無論、統合報告書等において特定の単語を多用している（ワードクラウドの文字サイズが大きい）からといって、銀行がその単語に関する具体的な取り組みを行っている証左にはならない。しかし、経営戦略において重視しているテーマであることには間違いなく、今後の経営の方向性を見通す上での有用な情報となる。さらに、本稿の手法には、業界全体における注力分野を統合報告書等の大量の文字データから効率的かつ素早く把握できるというメリットもある。

今回は、銀行業の統合報告書等を対象に分析を行ったが、不要な単語（ストップワード）の調整等を行うことにより、同様の分析を他業種に適用することも可能となる。さらに、統合報告書等以外の開示資料、例えば、中期経営計画書への応用も考えられる。

今後も重要性が高まっていく ESG 関連情報をはじめとする非財務情報を分析するにあたり、テキストマイニングや機械学習といった技術は有用な手段の一つになるだろう。