

論文 | Article

ESG エコシステムに向けた AI の貢献可能性

Potential Contribution of AI to ESG Ecosystem

越智 信仁

OCHI, Nobuhito

尚美学園大学
総合政策学部教授
Shobi University

2021 年 3 月

March.2021

ESG エコシステムに向けた AI の貢献可能性

Potential Contribution of AI to ESG Ecosystem

越智 信仁

OCHI, Nobuhito

[抄録]

比較可能な ESG (環境・社会・ガバナンス) 情報を起点として、資本市場のダイナミクスをドライビング・フォースに取り込めれば、ビジネスと持続可能な社会をつなぐエコシステムの循環に資する。こうした問題意識から本稿では、ESG に係る企業活動への市場規律向上を基本的視座とし、現下の ESG インデックスのばらつきによる比較可能性欠如という実務的課題の克服に向け、AI を活用した S-Ray による比較可能性改善の事例研究を通して、ESG エコシステムの実現に向けた AI の貢献可能性を展望する。

S-Ray は、メディア報道を含む膨大な外部情報源の迅速な解析に強みを有しており、ESG インデックスのばらつきを克服するユニークなツールとして、市場規律の向上に資するポテンシャルを有している。他方で、AI のブラックボックス問題や学習データ依存性に起因して、S-Ray を含め情報源・アルゴリズムの偏向等への疑念も現状では完全に払拭し切れていない。AI による ESG 評価の品質保証は難しい課題であるが、モデル策定プロセスの内部統制に係る検証等に関しては、監査保証論のフレームワークを基礎として役立てられる可能性があり、今後は会計・監査保証分野においても幅広い AI 研究コミュニティと緊密に連携した学際的な研究の深化が求められる。

キーワード

ESG エコシステム、AI (人工知能)、ESG インデックス、比較可能性、品質保証

[Abstract]

Incorporating the dynamics of capital markets into the driving force through comparable ESG (Environmental, Social and Governance) information, will contribute to the circulation of the ecosystem that connects business and a sustainable society. Based on this awareness of issues, this article focuses on improving market discipline for corporate activities by improving the comparability of ESG evaluations using AI in order to overcome the current practical issues of ESG index variability, and specifically starting with S-Ray's efforts utilizing AI, the author considers the potential contributions and challenges of AI toward the realization of an ESG ecosystem.

S-Ray has strengths in rapid analysis of vast numbers of external sources, including media coverage, and has the potential to contribute to improving market discipline as a unique tool for overcoming ESG index variability. On the other hand, due to the black box problem of AI and the dependency on learning data, doubts about the bias

of information sources and algorithms including S-Ray have not been completely dispelled at present. Quality assurance of ESG evaluation by AI will be a difficult issue in the future, but it may be useful based on the framework of auditing/assurance theory for verification related to internal control of the model formulation process, and it will be necessary to promote interdisciplinary research in close collaboration with a wide range of AI research communities.

Keywords:

ESG ecosystem, AI (Artificial Intelligence), ESG index, comparability, quality assurance

1. はじめに

ESG (Environmental, Social and Governance) 情報には S (社会) の領域を中心に規範的事項も多く含まれ、本来的に多様な特性を有している。このため、企業開示データの解釈における複雑性と比較困難性を内包し、これが ESG 評価機関の各種インデックスのばらつきを生む遠因ともなっている。しかし、新しい AI (人工知能) 技術を活用するなどして ESG 情報の透明性・比較可能性を高めるブレイクスルーが実現できれば、ESG 情報の開示を起点として、資本市場のダイナミクスをドライビング・フォースに取り込む形で、ビジネスと持続可能な社会をつなぐエコシステムの循環に資することが可能となる (World Economic Forum 2019, pp.6-7)。

こうした大局的な問題意識を背景に以下では、まず現下の具体的な実務的課題である ESG インデックスのばらつき問題の克服に向け、AI を用いた ESG 評価の比較可能性改善による企業活動への市場規律向上の可能性を指摘する¹。これを受けて次に、AI による ESG 評価の事例研究として、比較可能性の向上を意識した S-Ray にフォーカスし、その特徴をレビューしつつ ESG エコシステムの実現に向けた貢献可能性を展望する。そのうえで、こうした光の部分とは裏腹に影の部分として、AI のブラックボックス問題や学習データ依存性に起因して、ESG 評価の品質保証を巡る今後の課題や、信頼性向上に資する対応の方向性についても論及する。

2. ESG 情報の比較可能性による市場規律向上

2.1 ESG インデックスのばらつき問題

ESG 投資において、環境、社会、ガバナンスの多様な要素を、投資の分析、選択、管理に組み込むに際し、明確なリストや透明なフレームワークは今のところ存在しない。こうした中、年金等資産保有者から ESG 投資を条件に受託した運用機関は、自ら調査を行うだけでなく、外部の評価・格付機関の調査情報を利用することも多い。機関によって様々な

¹ 比較可能性に焦点を当てた ESG 評価の可能性と課題については、越智 (2019b) において AI 以外の分野を含めて包括的に論じている。

表記となっているが、業種特性も加味しつつ主に定性的な ESG 項目を基に最終的に定量的なインデックスを提供する点では概ね共通しており、いずれも投資家向けプロバイダーとして、ビジネスの視点で企業活動に与える重要な ESG リスク・機会への考慮が行われている。

しかし、現状の ESG インデックスは、評価データ・手法等の相違から同一対象に対しても機関によって評価のばらつきが大きいことが実証的に示されている (Chatterji et al. 2016, pp.1602-1607)。このため、企業内での管理情報等の公開の促進とともに、これを受けた評価機関によるデータベースの蓄積・拡充、ESG インデックスの精度向上に向けた評価手法の改善等が課題となっている (張替 2017, p.12, 20; 塩村 2017, p.41)。こうした中、従来の評価・格付機関による一般的な ESG 評価手法は、①企業が公表した情報を基に評価する方法、②企業への調査票に対して任意に回答した内容を基に評価する情報が主流であったが、近年では、これに③ AI を使った ESG 評価が登場してきている。

ビッグデータや AI の機械学習²によって、膨大かつ多様な定性的非財務情報の特徴の抽出・識別に役立つ可能性は積極的に評価されて然るべきであり、ESG 分野の情報検索において人間以上に情報認識の速度と精度が実現するとすれば、とりわけ ESG のうち S (社会) においてビッグデータと AI の活用可能性は高いとみられる。AI を活用した ESG 評価の収集・分析に関しては、既に少なからず商業化されており、実際に投資家が利用しているサービス提供機関として、次節で詳述する Arabesque S-Ray のほか、TruValue Labs、Deutsche Bank Research、三菱 UFJ トラスト投資工学研究所等が挙げられる。

このうち、2013年にサンフランシスコで設立された TruValue Labs は、メディア報道や NGO の発表、政府機関からの膨大な外部情報源のみからビッグデータを集積し、自然言語処理を用いた ESG スコアをタイムリーに算出しており、恣意性の懸念がある企業発表情報を一切用いないとしている点でユニークである (夫馬 2018)。また、従来の ESG インデックスでは同一企業の評価に対するばらつき、一貫性のなさが問題視される中であって (藤田 2018, p.5)、Arabesque (2013年設立) が 2017年に公表した AI による ESG 評価ツール (Arabesque S-Ray) が注目される。S-Ray の最大の特徴として、各種情報源の中に多くの ESG 評価機関が提供するインデックスも統合されており、比較可能性が担保された統一尺度となり得る点が挙げられる。

アラベスク社日本代表へのインタビュー記事である CSR コミュニケート (2017) によれば、創業者セリムは英バークレイズ銀行で ESG 投資戦略を立案するに際し、ESG 調査機関の評価に一貫性がなく (ばらつきが大きく) 利用を断念していたが、各機関が資金と労力を投入して下した評価結果には一定の価値を内包すると考えるに至り、AI を用いて異なる価値の最適な値を運用に役立てるべく 2013年に独立し、2017年の S-Ray 開発に至ったとされる。スコアリングに際しては、15言語 5万以上のニュースメディアから日々収集した情報に加え、上述したように複数の ESG 評価機関が提供する 200以上の ESG

² machine learning は、ISO/IEC 2382 において「機能単位が新しい知識・技能を獲得すること、又は既存の知識・技能を再構成することによって、自身の性能を向上させる過程」と定義されており (福田 2017, p.8)、AI を含む機械が課題を解決する能力を自ら学習していくために必要なメカニズムを解明することを目的とする研究分野である (荒屋 2004, p.105)。

指標をも組み合わせ独自の手法で企業評価を行っている (Kell 2017; 夫馬 2017)。

このほか、既存の ESG 評価機関の大手である Sustainalytics でも、日々の記事やレポート等を AI が自動検索し、重要な ESG 関連情報を収集の上、最終的な人間による判断に役立てているとされる (水口 2018)。近年では、2018 年 9 月に米国の Richmond Global Compass Capital が AI を活用した ESG ファンドを立ち上げたほか、2019 年 2 月にはドイツのインデックス開発会社 Solactive が、ESG 評価に AI を導入した新たなインデックスの開発を公表している。わが国でもサステナ社が、2018 年 1 月に ISO26000 を基礎として独自に開発した AI 評価システムで、上場企業を含む 4 千社以上の企業を対象に評価結果を検索できる企業評価サイトを公開するなど、AI を用いた ESG 評価の動きは国内外で今後とも強まっていくことが予想される。

2.2 ESG 情報の多様性と比較可能性

現状において ESG 評価・格付機関のインデックスは、信頼性のある横比較情報の提供という面では未だ発展途上といえる。情報の比較可能性には、同一企業内での時系列比較 (縦の比較可能性) とともに、企業間・業種間での比較可能性 (横の比較可能性) が存在するが、横の比較可能性が担保されることで、外部性問題を含めた企業活動に関して他社比較を通じた評判が生じ得る。このため、比較可能な非財務情報が増えることになれば、資本市場の評価機能による市場規律を強めることにもつながる。その際、ESG 情報で難しいのは、規範的事項を内包しているため、その内容が多義的で多様な情報特性を有する点である。

例えば³、モンサント社が遺伝子組み換え種子を導入したのは、持続可能な農業への貢献だろうか、あるいは人々の健康や生態系の健全性に対する脅威なのだろうか。ブリティッシュ・ペトロリアム社はグローバルな気候変動の問題の存在を認めたことで賞賛されるべきか、それとも化石燃料の開発を継続したことで非難されるべきか。ユニオンオイル社はビルマにおけるパイプライン建設プロジェクトで、労働条件を改善したので責任ある行動をしていると言えるのか、それとも抑圧的な軍事政権下の国で事業を継続したので無責任と言うべきなのだろうか。これらの問いに対し回答するには、個別具体的な状況に応じて多元的な判断を分権的に行わざるを得ない。人々の選好、信条、理想が互いに異なるという事実を前提にするならば、価値多元性の下で、どのようにして公共的意思決定に到達することができるかという問題でもある (越智 2015, p.67)。

結局、何が高潔な企業行動なのかについてはコンセンサスがないので、何を購入するか、誰のために喜んで働くか、どこに投資するか、といった社会的な選考を「投票する」行動を通じて、人々は自らの価値観を表現し、企業の慣行に影響を及ぼすことが望ましいと言える (ボーゲル/小松ほか 2007, pp.6-7)。そうした投資を行うに際し、投資家 (アナリスト) は、1つの企業をじっくり見るという切り口と、同業他社比較といった切り口の 2つの分析アプローチを有しており (黒田ほか 2017, p.23 「窪田発言」)、そもそも ESG 情報には多様で断片的な「モザイク情報」にとどまる場合も少なくないとみられるので、そうした情報の咀嚼に際しては、情報の比較可能性を高めることが投資家の相対的価値評価に資す

³ 以下の例示は、(ボーゲル/小松ほか 2007, pp.7-9) による。

ると考えられる(越智 2019a, p.35)。

例えば、国連の SDGs (持続可能な開発目標) は分野毎に整理された世界共通の課題であり、ともすれば広範・漠然としがちな外部性問題への対応をグローバルに比較可能な世界共通言語で各国・各経済主体の開示に取り込み得る点が、国際金融資本の注目を集めている所以であろう。資本市場での評価機能をドライビング・フォースとして活用するうえで、投資判断に影響を及ぼすような比較可能な非財務情報が有用であり、ESG 情報の多くが概括的あるいは定性的にしか把握できないとしても、同業他社対比でみて如何に社会・環境課題に取り組んでいるかの相対比較が可能となれば、投資判断の材料となり得る⁴。

そこでは、投資家にとっての情報の非対称性緩和とともに、開示主体に他社比較を意識させ開示を底上げする推進力としてもミニマムな共通開示基準の提示が望ましく(北川 2017, p.59)、かつ同様のビジネスモデルを持ち同様の方法で資源を使用する傾向がある同業種毎に、比較可能性を高める方向性が重視されよう。例えば、わが国でも、女性活躍推進法による女性の活躍に関する情報の開示義務化が、「MSCI 日本株女性活躍指数(WIN)」(2017年7月 GPIF 採用) というファンド組成を後押しするとともに、指数銘柄企業の多くはホームページ等へのシグナリングへとループしているほか、定期的に構成銘柄の入れ替えが行われることもあってジェンダー面で各社の取り組みを後押しする推進力ともなっている。また、GPIF (年金積立金管理運用独立行政法人) が ESG 投資のベンチマークとして 2018年9月に採用したカーボン・エフィシエント指数は、指数全体の炭素排出量を減らしたい投資家のニーズを満たすとともに、炭素効率情報の比較可能性を梃として、企業の炭素排出量に関する情報開示と関連情報の透明性の改善を促進する誘因としても機能している。

資本市場における市場規律を引き出すには、投資家等が各企業の非財務開示要求事項への対応を業界内横並びで評価可能な仕組みとすることが肝要であり、IFRS Foundation (2020) においても、資本市場において比較可能なサステナブル情報の重要性が強調されている。各社の固有情報を評価し個別銘柄のアルファを獲得する絶対的評価目線を具備し得ない場合等には、比較可能性を高めることが投資家の相対的価値評価に資するのであり、ここで比較可能な情報は、例えば低ボラティリティ銘柄戦略等と同様な手法で、リスク・リターン特性を識別するファクターとして利用することによって、システムティックリスク(いわゆるベータ)に見合った予想収益率に対する超過リターンを期待する投資手法に活用可能となる(越智 2019a, p.34)。

実際、GPIF (2019) の調査によれば、ジャッジメンタル運用では企業個社の深い理解に役立つ ESG 情報へのニーズが高い一方、システムティック/パッシブ運用等では定量化や比較可能性に優れた ESG 情報へのニーズが高いことが確認されている(同, p.49)。

⁴ 民間投資を呼び込むためには、グリーン・ウォッシュを排除し投資判断のための透明性の高い評価基準やツールが不可欠であり、そのためには ESG 開示項目をある程度収斂していく必要がある。その意味では IFRS Foundation (2020) や IFAC (2020) において、IFRS 財団の下に国際会計基準審議会(IASB) と並ぶ新たな国際サステナビリティ基準審議会 (ISSB) の創設を提案し、関連開示基準を統合する形で、比較可能性と信頼性を兼ね備えたグローバルな企業報告システムを展望していることは注目に値する。

こうした認識を踏まえ以下では、ESG 情報が本来有する多様な性格を踏まえつつも、比較可能な ESG 評価情報は ESG 投資の活性化を通じ市場規律の向上にも資するとの観点から、ESG 投資で用いられる ESG インデックスが比較可能な投資尺度になり得ていない現状の打開に向けて、比較可能な ESG 評価への貢献を標榜しているアラベスク S-Ray 社を採り上げ、同社の AI による ESG 評価の活用例をレビューした後、より一般的に AI の活用に残された課題についても考察する。

3. AI による ESG 評価の活用事例

3.1 アラベスク S-Ray 社のミッション

現在のアラベスク S-Ray 社の会長は、国連グローバル・コンパクト創設メンバー・元事務局長であったゲオルグ・ケル氏であるが、同氏が語る自社のミッションでは、「アラベスクは、企業と資本市場にサステナビリティ情報を提供するために、現在の金融に革命をもたらしている ESG と AI に基づいて構築され」「企業の ESG 評価の透明性の向上と AI をはじめとする最新のテクノロジーを通じたサステナビリティのメインストリーム化」を目指している（同社ホームページより）。なお、アラベスク S-Ray 社のボードメンバーは、会長以外にもコーポレート・サステナビリティ分野における世界的リーダー達が非常勤取締役を務めているのも特徴であり⁵、ESG の幅広い知見が、スコアリングの構成要素の組成等に際して役立てられている（同社日本支店代表の雨宮氏からのヒアリングに基づく）。

AI による ESG 評価機関として世界的に注目を集めるアラベスク S-Ray 社であるが、同社のこれまでの経緯を紐解くと、もともとは英国バークレイズ銀行の社内ベンチャーとして 2010 年頃から ESG クオオンツ運用に取り組みられてきた後、先述したように 2013 年に MBO でアラベスク・アセット・マネジメント社として分離独立し（創業者セリム）、2017 年には Arabesque S-Ray® を開発・リリースした。その後、アラベスク・アセット・マネジメント社からスピノフする形で、2018 年にはドイツに新しくアラベスク S-Ray 社（Allianz X、Commerz Real AG、DWS Group、Heleba Digital、Land Hessen が出資）、さらに 2019 年にはアラベスク AI 社（DWS Group が出資）が、それぞれ分離独立した経緯がある。現在は、アラベスク S-Ray 社⁶、アラベスク・アセット・マネジメント社⁷、アラベスク AI 社⁸ で企業グループを形成し、グループ全体で最先端の研究を利用して、持続可能で透明の高い金融ソリューションを提供していくことを目指している（日本支店は

⁵ バーバラ・クラムシーク（老舗の SRI 投資会社であるカルバート・インベストメンツ元 CEO 兼社長）、キャロリン・ウー（カトリック・リリーフ・サービス元 CEO 兼理事長）、ジョン・ラギー（ハーバード大学ケネディ・スクール教授〈人権・国際関係専門〉）、ヨランダ・カカバドス（世界自然保護基金（WWF International）元総裁）が名を連ねている。

⁶ 本文で詳しく後述するが、企業のサステナビリティに対する異なる側面からの評価（ESG、国連グローバル・コンパクト、地球温暖化への寄与、事業関与）、企業がサステナビリティ・データを管理することができるツールの提供や企業の報告書作成業務の支援等を行う。

⁷ 投資戦略の構築と投資商品の運用として、非財務データの統合と定量的投資プロセスの実施によって多様な運用手法（例えば、mutual funds, hedge funds, SMAs）を用い、世界各地に拠点（米国、欧州、アジア、中東）を有する。

2018年11月に設立)。

もともとアラベスク社は、先述したように、ESG 評価会社の評価軸間に相関性が存在しない課題へのチャレンジにから創業に至った経緯がある。すなわち、多数の ESG 評価機関から提供されるサステナビリティ関連データのばらつきという課題・問題点を踏まえ、それらを総合化し比較可能にすることを目的に開発されたプロダクト・ソリューションである。企業の公開情報と世界 170 か国にわたる 3 万以上の情報源から多くの ESG データを収集・分析しながら、世界約 80 か国、約 8 千社（うち日本企業は約 6 百社）の上場企業に対する日次評価を可能とした。そこでは、アナリストを配置せず人間の主観を排除し、AI による独自のスコアリング・メソッドロジーで ESG スコア等を算出しているところに特徴がある。

従来の ESG 評価を巡っては、①データの問題として、開示の量・質の不足や透明性の課題に加え、年次開示に限られた遅い開示という頻度の課題、②アナリストの問題として、定性判断が介在するため偏った情報を生み出す可能性があるほか、ESG は幅広いテーマを扱うだけに、温暖化については知見が高くても廃棄問題には知見が不足するなど、個人が扱うには対象が広すぎるという課題、③事例中心の CSR 報告書では、ポジティブ情報が中心で、ネガティブ情報は開示したがない傾向、さらにグリーン・ウォッシュという課題等が指摘されてきた。アラベスク社は、これら課題を解決し、頻度高く客観的な情報を提供するため、自己学習型定量モデルとデータを用いて世界の上場企業約 8 千のサステナビリティ・パフォーマンスを分析する定量データツールとして S-Ray を開発し、企業のサステナビリティ・パフォーマンスに毎日アクセス可能としたのである。

3.2 S-Ray の特徴とポテンシャル

S-Ray は AI 機械学習の力を活用して、人間判断に頼る ESG 投資ではなく、多様なデータ・ソースを体系的に組み合わせた技術主導のプラットフォームであり、リスクと機会の視点から企業にとっての重要性（マテリアリティ）を把握することで、ESG 投資のメインストリーム化に対応しようとしている。同社ホームページには、利害関係者として 4 つの用途（Investors, Corporates, Media, Public Use）が挙げられているが、ESG を投資に統合して超過収益を狙う ESG インテグレーションに役立てるため、基本的には投資家の視点でサービスが提供されており、それが他の用途も充足し得る関係性にあると理解される。

S-Ray の差別化要因としては、まず、ESG 自体が内包する多様性に鑑みて、複数のレンズを通して ESG の評価（写像）を試みており、企業の持続可能性の多面的な性質を反映するため、単純な単一の合計スコアの代わりに、複数の異なる持続可能性の測定値を提供している点が挙げられる。具体的には、4 つのスコアを通して多面的な評価が行われており、① ESG スコアでは企業経営を株価に影響を与える目線で評価が行われる一方で、②

⁸ できるだけ人間の主観（バイアス）を排除し客観的な運用につなげることを標榜し、ビッグデータと AI を活用することによって、アラベスク独自の AI エンジンを開発（独自のテクノロジーを使用して構築された約 2 億 6 千万のノードと約 7 億 5 千万のエッジで構成される大規模分散型計量グラフ）、AI エンジンによって生成されたシグナルを算出し投資商品に活用していく活動を展開している。

GC スコアにおいてはレピュテーション・リスクを主としてモニターしている。また、③選好フィルターは投資家各人の価値観に沿った投資行動をチェックするとともに、④気温スコアはパリ協定を意識しつつ気候リスクとその行動を評価している。

特徴の2つ目としては、日々のニュースを反映した日次スコアが、土日関係なく 365 日提供される点である。そこでは、3万超の NGO およびニュースシグナルを基に、約 1 億 5 千万のデータポイントを毎日更新することにより、約 8 千もの上場企業を網羅し、250 項目に及ぶ長期の ESG データが還元される。そして、そうしたデータが人間の主観を排除した定量的メソッドロジーに基づいて、行動バイアスを回避するために非アナリスト（人的）体制が敷かれている点も、特徴の3つ目として挙げられよう。

但し、データ項目の増加を継続的に行えるシステム・スケールの下、AI で多様なデータ・ソース（独自データのみならず他社契約によるデータ）からビッグデータを分析するアルゴリズム・アプローチが採られる一方で、スコアリングの構成要素としてグローバル・コンパクトなど多様な評価項目を採り上げるルール自体は人間が決めている点には留意が必要である。スコアリング項目については、ボードメンバーからのアドバイスを通じて改善に努めており、いったんルールが決まった後は AI 以外の判断余地を排除している由である（日本支店雨宮代表からのヒアリングに基づく）。

以上の特徴を有する S-Ray は、世界の上場企業約 8 千社の時価総額を 9 割以上カバーし⁹、増加し続ける多様なデータを取り込みながら、企業のパフォーマンスを評価する定量的測定基準を構築している。こうした S-Ray の定量的分析アプローチは、分析対象企業のカバレッジ拡大のみならず、より包括的かつ柔軟なスコアリングを継続的に可能としている。具体的には先述したように、企業のサステナビリティを評価するための透明化ツールとして、ESG パフォーマンスを 4 つの側面から分析したユニークなスコア等を用いており¹⁰、以下で諸指標の特徴を敷衍する。

① ESG スコア (Environmental Social Governance Score)

ESG スコア (0-100 点) は、株価との相関や株価へのインパクトを考慮したマテリアリティ・アプローチを採用しており、ESG データの中から企業の業績に直結するキーになるものがどれなのかを見つけ出し、ここに AI の機械学習を適用することで独自の分析から相関するファクターを導出する。株価に影響を及ぼす ESG 課題を分析するうえでは、セクター毎の特性を踏まえて重要性のウエイトには差異が設けられており、財務的重要性 (マテリアリティ) は、パフォーマンスへの寄与の高い ESG 項目をオーバーウエイトする

⁹ グローバルなカバレッジとしては、世界で 8 千社近い上場企業を網羅し、時価総額ベースで世界の 9 割超の企業のサステナビリティのパフォーマンスに顧客はアクセス可能となっている。地域別には、North America (約 3 千社)、Asia (約 2 千社)、Europe (約 2 千社) のほか、Australasia (約 4 百社)、South America (約 2 百社)、Africa (約 2 百社)、Middle East (約 1 百社) となっており、日本でも約 6 百社 (時価総額の約 9 割) がカバーされている。

¹⁰ こうした状況下、スコア向上に資する本来的な取り組みの改善が求められることは言うまでもないが、日々の AI 評価に対応し日本企業としても、より頻繁に、ESG の要約情報 (とりわけ時事的な S の情報) や数量的データ等のタイムリーかつきめ細かな Web 開示が求められよう。

機能を適用し、四半期毎にウエイトをリバランスしながら、長期的に優れたパフォーマンスを発揮できる持続可能な企業を特定していくことになる。

ESG スコアを算出する際、財務・株価に影響を与える ESG ファクターに焦点が当てられ、これを通じて事業活動で財務的に重要な ESG 課題について企業パフォーマンスを評価する。このため、企業の経営状態である長期的パフォーマンスを特定するアルゴリズムには、将来のリスク調整後のパフォーマンスに寄与する情報のみを使用されることになる。具体的には、(a) 22のサステナビリティ課題における企業のパフォーマンスを計測・比較し、(b) 相関性分析（マテリアリティ分析）を通じ、企業の財務パフォーマンスをより良く説明するサステナビリティ課題を特定する。(c) 結果として ESG 投資に際しマテリアリティを考慮することで、どの企業が長期的にアウトパフォームし、投資成果の向上が図れるかを分析することが可能となる。

② GC スコア (Global Compact Score)

GC スコア (0-100点) は、国連グローバル・コンパクト原則に基づいて、企業の社会的責任（行動規範）を評価する。あくまで企業の行動自体に焦点が当てられており、ESG スコアと異なり、業界の特異性や財務的重要性を考慮していない規範的な評価である。但し、GC スコアが社会的なインパクトを重視しているといっても、それは中長期のパッシブ運用（市場平均のベンチマークに連動する成果を目指す運用手法）を採る投資家を意識しているからであり、あくまで投資家志向であることに留意が必要である。すなわち、GC スコアが低い企業は、レピュテーション・リスクが高まることによって、公的なライセンスが失われる可能性などリスク要因で考慮すべき対象となるので、ESG スコアとともに、企業の持続可能性プロファイルの強力かつ包括的な理解に役立つのである。

国連グローバル・コンパクトは、世界 160 か国以上から 9 千を超える企業が署名している世界最大級の企業のサステナビリティ・イニシアチブであり、GC スコアは、国連グローバル・コンパクトの 4 つの中核分野である人権、労働権、環境、腐敗防止に基づいて企業の行動規範を評価する。先述したように、業種に関係なく GC 原則に対する企業行動を分析する規範評価であると同時に、企業のレピュテーション・リスクの指標として役立てられるが、4 つの中核分野におけるそれぞれ 1 つのスコアの裏に幾つかのサブスコアも連なっており、それらの集積によって企業の行動規範が評価される。そこでは一種の非補償型アプローチが採用されており、あるサブスコアのスコアが 50 を下回ると、そのサブスコアのウエイトを高める形で最終スコアに反映される。

③ プリファレンス・フィルター (Preferences Filter)

プリファレンス・フィルターは、企業の事業関与をチェックするサーチ・ツールであり、前の 2 つのスコアとはやや異質である。ビジネスにおいて懸念すべき活動（酒、タバコ、兵器、ギャンブル、遺伝子操作など）から得られる売上収入の程度を調査・チェック可能にすることによって、投資家の価値観である選好（Preferences）に適合している企業かどうかか濾過（Filter）できる。事業関与の有無は企業のスコアに影響せず、特定の事業関与の有無を知らせるだけで有無のシグナルの公表のみとなっているが、企業活動の透明性向上に寄与している。同フィルターは、あくまで ESG スコアおよび GC スコアを補完するも

のであり、ESG を投資プロセスに組み込むツールと位置付けられている。

一種のスクリーニングであり、アラベスク社もヨーロッパの伝統的な ESG 投資手法の流れを引き継いでいるともいえる。現在、約 22 千社（うち日本は約 3 百社）において企業ニーズに応じ確認したいときに活用され、顧客の価値観に基づいたポートフォリオの組成に役立てられている。S-Ray の設計思想には、サステナビリティの定義は 1 つではなく多様との理解が背景にあるため、人、企業、国によってサステナビリティの意味は異なり得ることに配慮されており、同フィルターでは企業の関心に応じて抽出の閾値を調整できる柔軟な手法が採用されている。すなわち、十数種類の懸念すべき活動分野において、それぞれ全体の売上に占めるウエイトが 5% 以上の場合にはフラグを立てることを基本としながらも、ユーザーのニーズ（価値観）に沿って、1% でも 20% でも柔軟にカスタマイズ可能な仕様となっている。

④ 気候スコア (Temperature Score)

気候スコアは、温室効果ガス (GHG) の排出を通じて、企業が地球温暖化にどの程度影響を与えているのかを数値化したもので、パリ協定のトピックを考慮に入れて 2019 年 9 月にリリースされた。実績の GHG 排出量を基礎に今後の気候変動への対応を評価したうえで、企業は地球温暖化にどの程度影響を与えているのかについて、2030 年や 2050 年を展望しベンチマークに収まっているかを数値で判断できるように、企業スコアは気温 (1.5°C、2°C、2.5°C など) の形で発表される。気温スコアは世界の約 3 千の上場企業、時価総額ベースでは 8 割近くをカバーしている。

そこでは、排出量を 2°C 水準までに削減する目標を有しているか (Target)、最近の排出量の変化は、今世紀半ばまでに排出量をネット・ゼロにし、地球の気温上昇を 1.5°C に抑制するために必要なトレンドを辿っているか (Trend)、スコープ 3 の排出量を情報公開しているか (Scope 3) などの視点が採り入れられている。但し、あくまで投資家が、将来の財務パフォーマンスの潜在的可能性をよりよく理解し、リスク調整後のリターンを向上させるために役立てられるため、ESG スコアと同様に業界のレンズを適用し、将来の財務実績を正しく予測可能な持続可能性基準のみが考慮に入れられることになる。

上に述べた S-Ray の定量テクノロジーにより、ユーザーはサステナビリティに関する付加的な影響を評価可能となる。3 か月前の S-Ray の企業スコアであれば、同社ホームページ上にエクセルデータのフォーマットで無償公開されているが、有料でより詳細な加工情報が土日関係なく 365 日更新されて入手可能となる¹¹。既に、世界の主要な金融機関、

¹¹ 提供サービスとしては、クライアントの独自要件・配信条件に合わせて約 8 千社の S-Ray スコア及びデータ (CSV ファイル形式) が毎日配信される (Enterprise License) ほか、例えば、日本株ファンドに組み入れられた ESG スコア、ウエイトの高い ESG 課題にどのように取り組んできたのかなど、時系列でみた重要性の分析やユーザーの運用ポートフォリオのサステナビリティ・パフォーマンス・データ (Portfolio Screening)、企業のサステナビリティ・パフォーマンスのスコア分析および競合分析、ある企業のある時点のスコアの寄与度分析や他社比較など (Company Snapshot)、企業のサステナビリティ・パフォーマンスに関する臨時レポート (Feature Overview) なども示される。

機関投資家、企業、コンサルティング会社が S-Ray を使用しており、公表・報道資料等によれば、海外では、スウェーデンの公的年金基金 API、ドイツ銀行、独インデックス開発大手 Solactive、アクセンチュア、BNY メロン、ステートストリート、S&P グローバル、BIMB インベストメント等が S-Ray のデータを使用している。また日本でも、第一生命が S-Ray データによる外国株式 ESG ファンドを組成しているほか、年金積立金管理運用独立行政法人 (GPIF) も自らの ESG 評価指標のデータに活用している。この間、アラベスク S-Ray 社は、QUICK 等とも提携し、ESG ソリューション事業の一環として S-Ray を提供している。

もともと S-Ray というネーミングは、健康診断のレントゲンである X-Ray にちなみ、S は Sustainability ということで、企業の持続可能性の健康診断を可視化する意図が込められている。その際、アラベスク模様と同じパターンが反復する一連の幾何学的美術形式であることにちなみ、日々、約 8 千の企業行動のパターンを比較可能な形で並べて、そこから優れた企業を見出す、すなわち横並びの評価を通じ、あるパターンから逸脱する企業をチェックする¹²。既存の ESG 調査機関が、一つの企業について定性面を含め専門アナリストが細かく分析した企業固有価値の探求に比較優位があるに対し、S-Ray では会社固有情報の細かな分析が十分できないとしても、競合他社との定量的なスコアで比較可能性を高め投資判断に役立つと同時に、自社の位置付けを横並びで把握可能にするという利用価値がある。こうした比較可能性を高める機能は、市場規律の向上を促進し、ビジネスと持続可能な社会をつなぐエコシステムの循環に資するポテンシャルを有していると評価できよう。

4. AI による ESG 評価の課題

4.1 ブラックボックス問題と学習データ依存性

上述した ESG 評価の分野に限らず、AI の機械学習技術は様々なシステムに埋め込まれ、実社会での応用・実用化が急速に広がっている。しかし、機械学習は決められた範囲内で適切な値を見つけ出すだけで例外に弱く柔軟性がない (松尾 2015, p.8) 中であって、機械学習技術そのものの研究開発が活発なのに比べて、それらの技術を組み込んだシステムの安全性・信頼性や品質保証に関する研究開発については、まだ取り組みが十分でない。そこにはグローバルに共通の課題があり、わが国でも科学技術振興機構研究開発戦略センターによる CRDS (2018a) (2018b) において、AI ソフトウェア工学の視点から、安全性・信頼性・品質保証に向けた方法論・技術体系の必要性が包括的に論じられている¹³。

深層学習は、人間が形式知化することが難しいレベルの判断規則まで自動獲得し得ると

¹² ディープラーニング (深層学習) によって、AI 自らデータ間の相関関係を分析し特徴量を自動的に見つけ出すことができるので、何かおかしいところに気づく異常検知というタスクは、高次の特徴量を生成できる深層学習の得意とするところである (松尾 2015, p.139, 155, 231)。

¹³ 国として適切な品質基準・安全性基準を定めることが求められ、その基準が十分な根拠を持ったものであることや、その基準を満たすようなシステム開発のための方法論が確立されていることが、国際競争力の源泉にもなるとされる (CRDS 2018b, p.12)。

ということが、人間を上回る精度の達成を可能にしている反面で、機械学習が判定したとき、その理由を人間が理解できないというブラックボックス問題が、システムの安全性を確保するうえで妨げとなっている。すなわち、機械学習技術は、いわば「前例に倣う AI」であり、多数の前例をデータとして与えると、その中に潜む規則性を見出し、新たに与えられたデータに対して、前例に倣った判断（データの分類・予測・異常検知等）ができるようになるのと裏腹に、人間にとって明示的な規則として理解することが困難なモデル（ブラックボックス）とならざるを得ず、このため出来上がった学習済みモデルにおける深層学習を人間が理解できないので、この解釈性が欠けているという問題が品質管理の課題を生むことになる（CRDS 2018a, p.5; 2018b, p.5）。

動作保証・品質保証等の考え方が整備されないまま、機械学習技術を組み込んだシステムの社会実装が急速に進むと、そこで発生した問題や事故が思わぬ社会問題化する懸念が指摘されている（CRDS 2018b, p.8）。まさに両刃の剣であり（Taylor 2016 p.4）、このため基礎研究として、機械学習自体の品質保証、全体システムとしての安全性確保、ブラックボックス問題への対策、問題を効率よく解く工学的な枠組み等が、重要な技術チャレンジとして浮上している。もとのアルゴリズムをきちんと実装できているのかとか、正しさの基準やその検証をどうするか、システム全体の品質をどう評価しどう管理するのか、何らかの問題が起きたときどうするかといった安全性基準の策定（および第三者評価の仕組み）という技術的な課題に加えて、制度・ガイドライン面の整備¹⁴も急務とされる（CRDS 2018a, p. ii, 6; 2018b, p.31）。

また、機械学習型システムは、従来のソフトウェア開発と異なり要件の厳密な記述やモジュール毎の検査ができず、学習用データが実データの統計的に適切なサンプルであることは保証されない¹⁵。こうした中、機械学習自体の品質はデータ集合が左右し、AIは臨機応変ということはできないため、訓練データに潜む差別・偏見をそのまま受け入れてしまう問題や、こうしたAIの脆弱性を悪用した攻撃も懸念される（Brundage et al. 2018, p.5）。実際に起きた問題事例として、Microsoftが開発した機械学習型チャットボット Tay が、2016年3月23日にインターネット上に公開されて、わずか一日で公開を停止するという事件があった。Tayは、悪意のあるユーザーとの対話から言葉を獲得した結果、差別と陰謀論に染まってしまう、不適切な発言を繰り返すようになってしまったのである（CRDS 2018b, p.8）。

悪意を持った攻撃だけでなく、設計が不十分なため意図しない有害な動作が発生する懸念（Amodei et al. 2016, pp.2-3）のほか、知らずに差別・偏見を取り込んでしまう問題もある。例えば、罪を犯した人間の再犯率を予測するために、人間による判断結果を用いて訓練したとすると、もしその判断を行った人間が特定人種に偏見を持っていたならば、機械学習による判断結果もその傾向を反映してしまう。また、個々の訓練データに偏見を持っ

¹⁴ 福田（2017）では、AI開発ガイドラインの必要性を論じる中で、機械学習技術に関わる知的財産権、プライバシー、セキュリティー等のガバナンス問題を採り上げている。

¹⁵ ランダムサンプリングして訓練データセットが出たと仮定し、ここから先は深層学習のアルゴリズムでモデルができるが、最初のステップがランダムサンプリングであるため、そもそも統計的機械学習は本質的に確率的とされる（CRDS 2018a, p.10）。

た判断があるか否かだけでなく、訓練データの分布の偏りや量が差別を生むケースも指摘されている。例えば、特定人種について誤認識が多いことが問題になった事例や、採用判定で性別による差別が起きてしまった事例等が報告されている (CRDS 2018b, p.9)。

さらに最近、社会的な面でも、深層学習における品質保証・動作保証の脆弱性の問題が指摘されている。例えば、CRDS (2018a) p.3によれば、車の自動運転の際に、標識の認識に深層学習を使う際、「止まれ」(STOP) の標識を認識できる深層学習モジュールに対して、「止まれ」標識に少しのノイズを乗せたものを見せて、速度制限の標識だと騙すことができたという報告があるとされる。事前の学習データとして、あらゆる場合を尽くすことは無理で、こういった攻撃ができてしまうので、動作や品質が保証できないならば、製造物責任 (PL 法) やリコール制度に該当しないか等も絡んできて、大きな社会問題になりかねない (CRDS 2018a, p.62)。こうした事項を含め、AI 技術の生産者・利用者双方のより一般的な法的・倫理的問題を調査する重要性も指摘されている¹⁶ (Future of Life Institute 2015, p.4; Russell et al. 2015, p.107)。

AI の社会実装 (事業化) においてリスク最小化のマネジメントが重要となることは、ESG 評価においても同様である。そこでは AI による判断や意思決定を行った結果として影響を受ける相手方に対し、アカウントビリティの履行が求められる (増子 2020, p.11)。ただ、人間の生命・身体に実害を与えない ESG 領域では、ミスの深刻性が相対的に低いともいえ、先述した工学的な課題を同様に抱えながらも、実用化サイドにおいて事業展開を比較的早く推し進める可能性がある。その場合には、機械学習型システムの品質管理基準として通常有すべき信頼性を判断するうえで、身体生命への危害と経済的不利益では、ダブルスタンダードの品質管理基準が首肯され得るか否かも含め、十分に議論を尽くす必要がある¹⁷。

ESG 評価の比較可能性向上により市場規律を促すにしても、AI の利用に関しては、先述したようにスコアリング・モデル (アルゴリズム) がブラックボックス化する中で、情報収集のアルゴリズムがバイアスを含み意図せざる偏向を生む懸念を現状では排除し切れない。また、高いスコアを得ようと近視眼的で形式的な企業行動を助長する場合には、短期的・機会主義的行動によって画一化が横行し、本来的に望ましい ESG 活動の多様性を歪めてしまう懸念もある (水口 2017; 2018)。AI による ESG 評価プロセス・結果への信

¹⁶ 久木田 (2017) では AI の倫理を議論する国内外の各種組織・会議等が紹介されており、日本は欧米に比べて遅れているとされるが (同, p.2)、国内外の報告書の多くは検討すべき項目を列挙している段階に過ぎない (西垣・河島 2019, p.20)。また、平野 (2019) では、ロボットを介した AI の負の側面に対しロボット法を発展させる必要性を論じる中で、行動規範としての国内外 AI 諸原則が紹介されている (同, pp.281-308)。

¹⁷ 例えば、SF 映画「2001 年宇宙の旅」で自らのシャットダウンを察知した AI (HAL) が、人間への攻撃を企てるシーンを記憶しているが、そのような事態を回避するため、必要に応じて AI システムを自動停止できるオフスイッチの設計も必要となるものの、このような状況は、少なくとも ESG 情報開示の評価分野では考えにくい。但し、同じく SF 映画で、全知全能のマザーマシン (AI) が統治する社会において、その判断に人間が盲目的・奴隷的に従うことを余儀なくされる危険性は、ESG 評価分野においても共通に存在しよう。

頼性付与は今後に残された課題であり、AIのみに頼り切ることの是非も論じる必要があるかもしれない。そもそも情報量の少ない中小企業等では、ビッグデータの検索というAIの強みが逆に弱点となり、機械学習では適切な評価ができないことになる。また、AI評価方法自体の信頼性ととも、情報源となる企業開示情報、各種メディアでのニュース報道等に多くを依存することの懸念も残り、チェリー・ピッキングがないとしても、開示企業のESG / SDGs ウォッシュや特定メディアの偏向報道、誘導記事等を受容し無批判に受け入れて評価を下してしまう危うさを内包していよう (越智 2019b, p.26)。

機械は与えられた問題を解くことはできても、問うべき問題を特定することはまだできないので、ブラックボックス型モデルを受容し無批判に受け入れるのではなく、その有効性をクオンツ分析者自身が総合的に判断を下す必要がある (伊藤 2019, p.48)、規範的事項のルールメイキングも含め人間が関与するプロセスは不可欠と考えられる。最終的には人間とAIの判断バランスの線引きが重要になり、そこではブラックボックスであるAIの判断の責任は誰がどのように負うべきか、AIと人間が共存する社会¹⁸のありようが問われることになる (瀧 2020a, p.2; Soares 2016, p.6)。

4.2 信頼性・品質保証に向けた方向性

ブラックボックス問題に伴うAIシステムの解釈可能性の課題については、適切に設計された評価アプローチが確立して初めて、AIシステムによる決定の質の評価が可能になる (Doshi-Velez and Kim 2017, p.9)。AIによる適正性の判断が社会に信頼されるには第三者の検証が必要であり (小俣 2020, p.204)、ここでは準拠基準と検証基準の必要性という意味で、一種の情報の監査可能性と同様のフレームワークが求められよう。AIによるAIの監査可能性に関しては小俣 (2020) p.204でも論及されているが、データをAIが取得しデータの信頼性の判断までAIが行う場合にも、データセットの幅や複数のデータセットによる判断差異の識別を通して、人間が規範的に判断する余地が残るかもしれない。

Google DeepMindでは、自社のAI開発ガイドラインを、①仕様 (AIシステムが設計者の真の望みに従って行動するよう動機づけられるようにすること)、②頑健性 (予期しない敵対的攻撃を含めデータの摂動<小さな攪乱・ずれ>に耐えるようにシステムを設計すること)、③保証 (システムアクティビティの監視と制御) という3つの視点からまとめている。ここでは、慎重に検討し計画を立てることで、類似の問題や脆弱性を組み込むことを回避できるよう、技術的なAIの安全性に関する体系的な計画に向けた有用な問題のフレーミングの確立が求められており、将来のAIシステムが「うまくいけば安全」であるだけでなく、堅牢で検証可能に安全であることを保証することが求められている (Ortega et al. 2018)。

¹⁸ 例えば、医療AIシステムの診断がどのようにして結論に達したかの説明とともに診断を下すの際に、医師が承認前に推論プロセスを検査可能なフレームワークが構築されている (De Fauw et al. 2018, p.1348)。このように人間では気づかないようなエラーのパターンを認識するようなAIシステムが、人間が見過ごすようなエラーの存在を示すことによって、人間の判断を補完し得る関係性は一つの参考となるのではなかろうか。なお、AI化の下でも医師等の業務として、(患者の気持ちを) 分析、(手術を) 説得、責任 (をとる) など、人とのコミュニケーションの分野は残ることになる (藤岡 2020, p.239)。

品質保証の問題は ESG 評価分野に限った話ではないが、AI プロダクトの品質保証に関する調査・体系化、適用支援・応用、研究開発を推進するとともに、AI プロダクトの品質に対する適切な理解を啓発する活動として、2018年には「AI プロダクト品質保証コンソーシアム QA4AI (Consortium of Quality Assurance for Artificial-Intelligence-based products and services)」が産学 27 の発起人・団体で設立されている。これに先立ち 2014 年から毎年開催されている FAT/ML (Fairness, Accountability, and Transparency in Machine Learning) ワークショップ等では、機械学習における公平性・説明責任・透明性等が議論されるようになってきている (CRDS 2018b, pp.35-36)。

他方で、置かれた状況は異なるにせよ、過去の(人間による) ESG 評価・格付インデックスを巡っても同様にブラックボックス問題は発生しており、その対応を巡る経験が、AI による ESG 評価への対応を考えるうえでも参考になるのではなかろうか。すなわち、既存 ESG 評価・格付機関は大小含め数多く設立されてきており、乱立する評価・格付商品の透明性・品質を高める目的で、2011年に GISR (Global Initiative for Sustainability Ratings) が設立され、2013年には ESG 格付に求める原則等として「持続可能性格付基準 (Sustainability Ratings Standard)」も策定された¹⁹。

また、ESG 評価・格付機関の連合体である ARISE (Association for Responsible Investment Services) では、2003年に「責任投資調査に関する自主品質基準 (ARISTA 3.0)」を策定した後も定期的に改訂を重ねており、そこでは「品質管理原則」を掲げるとともに、ARISE の認証委員会が監査人とも連携しつつ機関認証を行う仕組みも採り入れている²⁰。同様に、今後とも AI による ESG 評価の実装化が進展し、一団の業界が形成されるようになれば、AI 評価ツールの信頼性、品質保証を高めるため、業界内自主団体の組織化とともに、AI クオオンツとしての職業倫理規範・行動原則の策定、デュープロセス認証の枠組みといった周辺の制度インフラの整備に取り組む必要があるだろう。

ブラックボックス化に伴う、意図した恣意性にせよ、意図しないバイアスにせよ、これらを排除するためには、①訓練データ・テストデータの質と量、十分性と適切性、② AI による判断の合理性の評価が可能とならなければならない (瀧 2020b, pp.181-182)。そこでの品質保証として本来求められるのは、結果としての「プロダクト品質」であることは言うまでもない。しかし、そのアルゴリズムがオープンにならなければ、学習済みの製品・商品だけを調べても、リバースエンジニアリングで分解したり動作を解析したりして仕様や仕組みを明らかにすることができない以上、学習アルゴリズムを秘匿したままの学習済みの製品・商品の学習結果から、学習アルゴリズムを推定するのは不可能である (松尾 2015, p.246)。

こうした難点を克服するため「説明可能な AI (Explainable Artificial Intelligence ; XAI)」に向けた議論もなされているが²¹、AI の内実がブラックボックスのまま学習ア

¹⁹ ただ、評価基準・プロセスの詳細は公開されていない中であって、中立的な第三者によるインデックスへの信頼性付与という面では課題を残している。

²⁰ ただ、大手機関を巻き込んだイニシアティブとなり得ていない (水口 2015)。

²¹ 原 (2019) によれば、XAI はアメリカの国防高等研究計画局 (DARPA) が主導しているプロジェクトであり、機械学習モデルの出力に加えて、その出力を補助する追加の情報 (モデルの解釈、判断根拠の説明など) を出力する技術一般及び研究分野全体を指す用語とされる。

ルゴリズムをきちんと捕捉できない（説明不能な）場合²²には、次善の策として、その品質の測り方を「プロセス品質」の保証に切り替える必要もあるかもしれない。システム工学の視点からの AI 品質保証（安全性）の分析は筆者の能力を超え監査保証論の射程からも外れるが、モデル策定プロセスの内部統制に係る検証等に関しては、監査保証論のフレームワークが基礎として役立てられる可能性がある。AI による ESG 評価の信頼性確保に向けた考察は、引き続き難しい今後の課題ながら、AI による ESG 評価という非財務情報の策定プロセス（内部統制）の適切な構築と運用については、監査保証の実務や研究が貢献できる余地はあるのではなかろうか²³。

5. おわりに

AI を活用した S-Ray は、ESG インデックスのばらつき問題を克服するユニークなツールであり、比較可能性の改善による市場規律の向上に資するポテンシャルを有している。他方で、一般に AI はメディア報道を含む膨大な外部情報源の迅速な解析に強みを有する反面で、S-Ray を含め情報源・アルゴリズムの偏向等への疑念も現状では完全に払拭し切れておらず、今後は評価の透明性を増しながら、各種市場で利用可能な指標として発展していくことが求められる。その際、現状の AI 活用の ESG 評価は、資本市場における商業ベースでのサービス提供として発展してきているが、今後は社会的価値の実現にも一層循環し得るよう、より幅広いステークホルダーからの評価も交えた多面的な切り口による比較可能な ESG 指標に向け、政府や NGO 等非営利組織による AI 活用の進展も期待される²⁴。

AI を活用した ESG 評価の多くは、先述した GC スコアの規範的評価事例を除けば、主として企業活動への財務的重要性に焦点を合わせている。このため、現下の資本市場における ESG 投資という枠組みの下で、財務報告の重要性（マテリアリティ）の考え方において、人間に対して害を及ぼすリスクは、ビジネスにとってのリスクと機会を通して考慮されることになる。具体的には、GRI（Global Reporting Initiative）の重要性マトリックスの縦軸（ステークホルダーの評価や意思決定に対して実質的な影響を与える項目）は企業の多くが採用している一方で、横軸（報告組織が経済、環境、社会に与える著しい影響＜インパクト＞を反映している項目）については、投資家向けの財務報告では自社にとって

²² この結果、データをたくさん持っている企業が高いレベルの深層学習の技術も手に入れられると、他の企業もそこにデータを集めざるを得ず、いったん差がつくと逆転するのは困難となり、パソコンの OS をマイクロソフトに、CPU をインテルに握られたのと同様に、少数のプレーヤーが市場を席捲する現象が AI の分野でも生じ得る（松尾 2015, p.247）。ただ、AI による ESG 評価の場合には、その多くが公表された情報を基礎としているため通常は想定し難いが、アンケート等の非公表データの比重が高まる場合には、ESG 評価分野においてもデータ寡占問題は生じ得よう。

²³ 瀧（2020c）p.257 においても、社会におけるテクノロジーの適切な運用に監査の実務や研究が貢献することの期待を、より一般的な文脈で述べている。

²⁴ 例えば、TCFD（2018）では、AI を用いて 1,734 社の財務報告者や統合報告、サステナビリティ報告書等をトレースし、TCFD（2017）の浸透状況（中核的要素であるガバナンス、戦略、リスク管理、指標と目標の開示状況）を、8 つのセクター毎、企業規模別、地域別のマトリックスにして確認している。

のインパクト（ビジネスのリスクや機会からみた重要性）に置き換えられる傾向にある（越智 2018, p.73）。

ただ、人間の目線からはビジネスの論理を超えた人間にとっての重要性に別途の考慮が必要になる。国連の SDGs（Sustainable Development Goals）では「誰も置き去りにすることなく、すべての人にとって尊厳ある生活を現実のものとするため」、グローバルな諸課題を設定し、2030年を目標年に国連加盟国とともに幅広い経済主体（自治体、企業、NGO 等）による総合的な取り組みを推進しており、そこでは自然資本や貧困と格差、食料、健康、教育、ジェンダーなど、人間もしくは人間が属するコミュニティが安全に生存するうえでの基底価値も対象となっている。企業が関係した人間への脅威（リスク）である外部不経済として、環境破壊や地域衰退、貧困の連鎖など分野横断的に人間の直面する脅威が多様化する中で、各企業のリスクマネジメントを超える公共的な要素への視角も求められるようになってきている。

ビジネスではなく人間にとっての重要性は、本来的に SDGs や GRI が目指すマルチステークホルダーへの考慮において達成されるべき事柄となるが、そうした分野においても従来のレポーティングに依存した情報収集や評価の限界を克服すべく、AI を活用した評価・分析フレームワークが構築できないであろうか。そこでも企業の作為・不作為の見える化に向けた横断的で比較可能な情報提供を通じ、当初は財務に関係なかった事象も、資本市場のみならず商品・労働市場等における評判が介在することで、正負のインタンジブルズとして企業価値の構成要素にも循環する。このため、評判を意識したマネジメントを企業から引き出し得るので、そうした事象は投資家の ESG 投資の考慮においても無縁ではなくなり、結果的にマネジメント対応のインセンティブを企業から引き出すことが可能になると考えられる（越智 2019b, pp.27-28）。

こうした評判による機能を作動させるに際し、今後は AI による ESG 評価が資本市場のみならず商品・労働市場等での存在感を増し、より広いステークホルダーにも利用可能な指標として活用が進むことが期待される。ビジネスのリスク・機会に直接関係しなくても社会に規範選択を問う重要事象に関し、企業の取り組みを比較可能な情報として提供する意義は大きい。そこでも評価の透明性・信頼性を増す制度インフラが非常に重要となるだけに、ESG 評価分野における AI 活用の品質保証に資するよう、会計・監査保証分野においても幅広い AI 研究コミュニティと緊密に連携した学際的な研究の深化が求められる²⁵。

【参考文献】

Amodei, Dario., Chris Olah, Jacob Steinhardt, Paul Christiano, John Schulman, and Dan Mané (2016) Concrete Problems in AI Safety, arXiv:1606.06565 [cs.AI].

Arabesque S-Ray website, <https://www.arabesque.com/s-ray/> (2021年1月5日閲覧).

Overview, <https://arabesque.com/docs/sray/Introducing%20Arabesque%20S-Ray.pdf>

²⁵ ICAEW（2019）では、会計プロフェッションの視点から、AI 利用に際しての倫理と説明責任の問題を論じているほか、日本監査研究学会の課題別研究部会（主査：瀧博）でも、AI 等の利用に絡んだテーマが取り上げられているが、AI による情報の品質保証に向けた学際的検討の深化は今後の課題に残されている。

- Methodology (v.2.6) September 2020, <http://arabesque.com/docs/sray/S-Ray%20Methodology%20v260.pdf>
- Brundage, Miles et al. (2018) The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation, arXiv:1802.07228 [cs.AI].
- Chatterji, Aaron., Rodolphe Durand, David I. Levine, and Samuel Touboul (2016) Do Ratings of Firms Converge?: Implications for Managers, Investors and Strategy Researchers, *Strategic Management Journal* 37(8), pp.1597-1614.
- De Fauw, J., Ledsam, J.R., Romera-Paredes, B. et al. (2018) Clinically applicable deep learning for diagnosis and referral in retinal disease, *Nature Medicine* 24, pp.1342-1350, DOI: <https://doi.org/10.1038/s41591-018-0107-6>
- Doshi-Velez, Finale., and Been Kim (2017) Towards A Rigorous Science of Interpretable Machine Learning, arXiv:1702.08608 [stat.ML].
- Future of Life Institute (2016) A Survey of Research Questions for Robust and Beneficial AI, https://futureoflife.org/data/documents/research_survey.pdf (2021年1月5日閱覽).
- Institute of Chartered Accountants in England and Wales (ICAEW) (2019) New Technologies, Ethics and Accountability, <https://www.icaew.com/-/media/corporate/files/technical/technology/thought-leadership/new-technologies-ethics-and-accountability.ashx> (2021年1月5日閱覽).
- International Federation of Accountants (IFAC) (2020) Enhancing Corporate Reporting: The Way Forward, <https://www.ifac.org/knowledge-gateway/contributing-global-economy/discussion/enhancing-corporate-reporting-way-forward> (2021年1月5日閱覽).
- IFRS Foundation (2020) Consultation Paper on Sustainability Reporting, <https://cdn.ifrs.org/-/media/project/sustainability-reporting/consultation-paper-on-sustainability-reporting.pdf> (2021年1月5日閱覽).
- Kell, Georg (2017) S-Ray: The Power of Sustainability Information, <https://www.eco-business.com/opinion/s-ray-the-power-of-sustainability-information/> (2021年1月5日閱覽).
- Ortega, Pedro., Vishal Maini and the DeepMind safety team (2018) Building Safe Artificial Intelligence: Specification, Robustness, and Assurance, DeepMind Safety Research, <https://medium.com/@deepmindsafetyresearch/building-safe-artificial-intelligence-52f5f75058f1> (2021年1月5日閱覽).
- Russell, Stuart., Daniel Dewey, and Max Tegmark (2015) Research Priorities for Robust and Beneficial Artificial Intelligence, AI Magazine, Winter, Association for the Advancement of Artificial Intelligence, <https://doi.org/10.1609/aimag.v36i4.2577>.
- Soares, Nate (2016) The Value Learning Problem, In: Ethics for Artificial Intelligence Workshop at 25th International Joint Conference on Artificial Intelligence, <https://intelligence.org/files/ValueLearningProblem.pdf> (2021年1月5日閱覽).

- Taylor, Jessica., Eliezer Yudkowsky, Patrick LaVictoire, and Andrew Critch (2016) Alignment for Advanced Machine Learning Systems, MIRI, <https://intelligence.org/files/AlignmentMachineLearning.pdf> (2021年1月5日閲覧).
- TCFD (2018) Status Report, <https://assets.bbhub.io/company/sites/60/2020/10/FINAL-2018-TCFD-Status-Report-092518.pdf> (2021年2月18日閲覧).
- World Economic Forum (2019) Seeking Return on ESG: Advancing the Reporting Ecosystem to Unlock Impact for Business and Society, http://www3.weforum.org/docs/WEF_ESG_Report_digital_pages.pdf (2021年1月5日閲覧).
- 荒屋真二(2004)『人工知能概論—コンピュータ知能から Web 知能まで【第2版】』共立出版.
- 伊藤敬介 (2019) 「AI、ビッグデータとクオンツ運用」『証券アナリストジャーナル』57巻4号、pp.44-49.
- 越智信仁 (2015) 『持続可能性とイノベーションの統合報告—非財務情報開示のダイナミクスと信頼性』日本評論社.
- 越智信仁 (2018) 『社会的共通資本の外部性制御と情報開示—統合報告・認証・監査のインセンティブ分析』日本評論社.
- 越智信仁 (2019a) 「財務・非財務情報の弾力性と比較可能性を巡る考察」『経営分析研究』35号、pp.27-40.
- 越智信仁(2019b)「比較可能な ESG 評価の可能性と課題」『インベスター・リレーションズ』13巻1号、pp.17-31.
- 小俣光文 (2020) 「AI 時代に必要とされる監査人の能力」瀧博編著『テクノロジーの進化と監査—AI とデジタル技術が拓く新たな監査の可能性』同文館出版、pp.189-205.
- 北川哲雄 (2017) 「フェア・ディスクロージャー規制とアナリスト」北川哲雄編著『ガバナンス革命の新たなロードマップ—2つのコードの高度化による企業価値向上の実現』東洋経済新報社、pp.43-61.
- 久木田水生 (2017) 「人工知能の倫理—何が問題なのか」総務省 AI ネットワーク社会推進会議平成 23 年 11 月 6 日資料、https://www.soumu.go.jp/main_content/000520384.pdf (2021年1月5日閲覧).
- 黒田康平・高畑修一・窪田真之・水口啓子・男澤江利子・川西安喜 (2017) 「特集 座談会 ディスクロージャーを巡る諸課題」『季刊会計基準』59号、pp.6-24.
- 国立研究開発法人科学技術振興機構 (JST) 研究開発戦略センター (CRDS) (2018a) 「俯瞰ワークショップ報告書 機械学習型システム開発のパラダイム転換」, CRDS-FY2017-WR-11, <https://www.jst.go.jp/crds/pdf/2017/WR/CRDS-FY2017-WR-11.pdf> (2021年1月5日閲覧).
- 国立研究開発法人科学技術振興機構 (JST) 研究開発戦略センター (CRDS) (2018b) 「(戦略プロポーザル) AI 応用システムの安全性・信頼性を確保する新世代ソフトウェア工学の確立」, CRDS-FY2018-SP-03, <https://www.jst.go.jp/crds/pdf/2018/SP/CRDS-FY2018-SP-03.pdf> (2021年1月5日閲覧).
- CSR コミュニケート (2017) 「ESG 情報 開示の最前線『S-Ray』」こちら CSR 革新室 12 月 19 日、<http://www.csr-communicate.com/csrinnovation/20171219/csr-32287> (2021年1月5日閲覧).

- 塩村賢史 (2017) 「GPIF の ESG 投資の取組みと今後の展望 (特集 政策が市場・運用を変える)」『月刊資本市場』387号, pp.38-45.
- GPIF (2019) 「GPIF 委託調査研究 (ニッセイアセットマネジメント): ESG に関する情報開示についての調査研究 報告書【概要】」, https://www.gpif.go.jp/investment/research_2019.pdf (2021年1月5日閲覧).
- 瀧博 (2020a) 「序論—テクノロジーの進化と監査における研究課題」瀧博編著『テクノロジーの進化と監査—AI とデジタル技術が拓く新たな監査の可能性』同文館出版, pp.1-9.
- 瀧博 (2020b) 「AI の諸特性と監査制度への示唆」瀧博編著『テクノロジーの進化と監査—AI とデジタル技術が拓く新たな監査の可能性』同文館出版, pp.173-187.
- 瀧博 (2020c) 「結論—総括と今後の課題」瀧博編著『テクノロジーの進化と監査—AI とデジタル技術が拓く新たな監査の可能性』同文館出版, pp.248-258.
- 西垣通・河島茂生 (2019) 『AI 倫理—人工知能は「責任」をとれるのか』中央公論社.
- 原聡 (2019) 「私のブックマーク 説明可能 AI (Explainable AI)」『人工知能』34巻4号, pp.577-582.
- 張替一彰 (2017) 「エンゲージメントと ERM の視点から見た ESG のあり方—ROIC/WACC 経営と持続的成長をキーワードとして」日本価値創造 ERM 学会 2017年度第3回セミナー配布資料.
- 平野晋 (2019) 『ロボット法—AI とヒトの共生にむけて【増補版】』弘文堂.
- 福田雅樹 (2017) 「『AI ネットワーク化』およびそのガバナンス」福田雅樹・林秀弥・成原慧編著『AI がつなげる社会—AI ネットワーク時代の法・政策』弘文堂, pp.2-44.
- 藤岡英治 (2020) 「医療分野におけるテクノロジーの進化から監査への示唆」瀧博編著『テクノロジーの進化と監査—AI とデジタル技術が拓く新たな監査の可能性』同文館出版, pp.221-242.
- 藤田香 (2018) 「SNS や NGO の評価も取り込む AI を活用した ESG 評価が始まる」『日経 ESG』11月号, pp.4-5.
- 夫馬賢治 (2017) 「アラベスク・アセット・マネジメントの企業サステナビリティ分析ツール『S-Ray』—部門長が語る開発背景と将来展望」Sustainable Japan, 9月19日, <https://sustainablejapan.jp/2017/09/19/arabesque-s-ray/28148> (2021年1月5日閲覧).
- 夫馬賢治 (2018) 「米 TruValue Labs、人工知能活用の画期的な企業分析ツール」Sustainable Japan, 1月19日, <https://sustainablejapan.jp/2018/01/19/truvalue-labs-insight-360/30154> (2021年1月5日閲覧).
- ボーゲル, デービッド/小松由紀子・村上美智子・田村勝省訳 (2007) 『企業の社会的責任 (CSR) の徹底研究: 利益の追求と美德のバランス—その事例による検証』一灯舎.
- 増子和起 (2020) 「AI の利用とアカウンタビリティ」日本社会関連会計学会スタディグループ最終報告書『SDGs 時代の社会的評価指標開発への挑戦』, pp.8-12.
- 松尾豊 (2015) 『人工知能は人間を超えるか—ディープラーニングの先にあるもの』KADOKAWA.
- 水口剛 (2015) 「ESG 評価を評価する—ESG 評価の信頼性確保に向けた動き QuickESG

研究所 6月 5日 , <https://www.esg.quick.co.jp/research/54> (2021年 1月 5日閲覧).
水口剛 (2017) 「AI と ESG 評価」QuickESG 研究所 6月 12日 , <https://www.esg.quick.co.jp/research/513> (2021年 1月 5日閲覧).
水口剛 (2018) 「テクノロジー産業と ESG 投資—その光と影」QuickESG 研究所 12月 11日 , <https://www.esg.quick.co.jp/research/965> (2021年 1月 5日閲覧).

(謝辞) Arabesque S-Ray 社日本支店の雨宮寛代表には、同社に関する情報収集や事実確認等に御協力いただいたことを、この場を借りて厚く御礼申し上げます。

