

資産運用におけるオルタナティブ・データ活用の可能性と課題

伊藤 健(野村証券金融工学研究センター)、佐藤 広大(野村総合研究所)、
(監修) 関 雄太

■ 要 約 ■

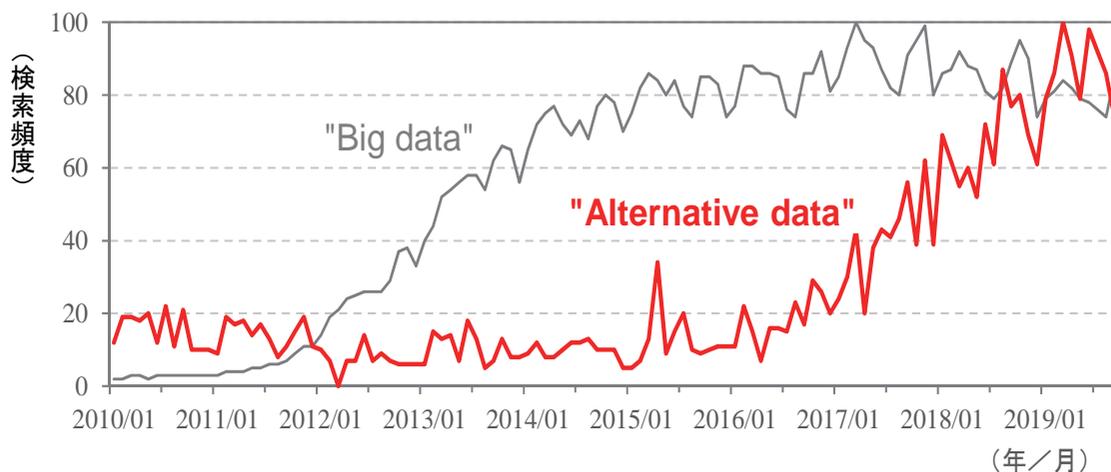
1. 米国大手クオンツ・ヘッジファンドを中心に活用が進んできたオルタナティブ・データは、近年、伝統的な資産運用会社やプライベート・エクイティの間でも注目度が高まっている。背景として、大量データの処理技術や人工知能技術等の汎用化、Mifid II など規制の改革、アルファ戦略の差別化が必要とされる運用環境の3つが挙げられる。
2. オルタナティブ・データは、伝統的なデータを補完することで運用者の意思決定に資するものであるが、即時性、詳細性、固有性の3点で優れていると考えられる。また、企業業績等の定量情報をより迅速・詳細に測定したり、企業の評判や ESG に関する取り組みなど従来は定性的にしか評価できなかった事項を定量的に測定できる可能性を持つと考えられる。
3. 資産運用会社がオルタナティブ・データを利用しようとする際、多種多様なデータやベンダーを逐一検証し、自らの運用目的に適した有用な情報を取り出すためには多くの時間と資本が必要である。
4. 利用者層の拡大に伴い、オルタナティブ・データの供給サイドにも大きな変化が起きている。新興ベンダー同士の合併・買収に加え、伝統的な大手情報ブローカーの参入、取引所や格付け機関によるデータ事業の拡充、プライベート・エクイティ・ファームによるデータ部門の取引、セルサイド金融機関による新規データ事業の模索が代表的な動きであり、今後が注目される。
5. 今後も、他者が目をつけていないデータや独創的な活用方法を探索する競争は続くと考えられる。適切なパートナーとの協業により、オルタナティブ・データを活用することで資産運用会社としての自社の強みを強化・補完していくことが、重要となるのではないだろうか。

I 資産運用における非伝統的データの活用

1. ビッグデータからオルタナティブ・データへ

今日、データの集積と効率的な活用が企業の競争力を左右すると言われ、各企業は市場や顧客に関するデータを分析し、市場動向の把握や、自社製品・サービスの改善に活かすことが求められている。資産運用においても、多種多様なデータから新しい知見を取り出し、投資判断に活かすための取り組みが数多く報告されている¹。これらの新しい形態のデータは、株価、財務、経済統計等の伝統的な金融データと対比する形で、オルタナティブ（代替）・データと総称され、伝統的データに加えてオルタナティブ・データを用いることで、企業や経済の状態をより迅速に、詳細に、多角的に評価できると考えられている²。なお、膨大な量のデータから知見を抽出するという観点では以前からビッグデータの活用という戦略も存在する。ビッグデータがデータの大きさに注目した呼称である一方、「オルタナティブ・データ」は、データサイズに関わらず、資産運用に応用され得る非伝統的データ全般を指す。グーグル検索動向に基づけば、「オルタナティブ・データ」という語句への注目は直近数年間で高まり、「ビッグデータ」への注目度を凌駕しつつある（図表 1）。

図表 1 「オルタナティブ・データ」の検索動向（月次、2010年1月～2019年9月）



(注) 縦軸の検索頻度は対象期間中の相対的スコア。対象地域は世界全体。
 (出所) Google Trends から野村証券金融工学研究センター、野村総合研究所作成

¹ 例として、下記の文献および文献中の参考文献を参照。
 饗場行洋、山本裕樹「データサイエンスと新しい金融工学」『財界観測』2018年春号。
 西家宏典、津田博史「従業員ロコミを用いた企業の組織文化と業績パフォーマンスの関係」『証券アナリストジャーナル』2018年7月号。
² 嶋村武史「資産運用におけるオルタナティブ・データの活用と日本への示唆」『NRI 知的資産創造』2019年3月号。

代表的なオルタナティブ・データとして、自然言語処理 AI や画像認識 AI を用いてテキスト（単語、文章、数字などで構成される文字列情報）・画像等の非構造化データをスコア化して得られるセンチメント・データ³や、各種統計データが挙げられる。AI を用いて自動的にビッグデータを処理することで、人間が処理するよりも迅速に、また場合によってはより詳細な情報を取り出せることが期待できる⁴。得られたオルタナティブ・データは、従来データを補完もしくは代替し、既存の運用戦略を洗練させるために利用される。

オルタナティブ・データの一例として、貨物の出入港を示す通関データが挙げられる。米国の港を通して輸出入する商品については船荷証券（Bills of Lading）の提出が義務付けられており、貨物の概要および輸入元（輸出先）の企業名および港等が記載されている。大量の船荷証券データを分析することで、国別の貿易額の増減、業種別の取引量、また個別企業の生産・販売動向について示唆を得られると考えられる。例えば自動車メーカーが鋼材の輸入を増やしたとすれば、将来的な生産量増加の可能性が見込まれるといった具合である⁵。船荷証券そのものは従来から存在するデータであるが、基本的に紙に印刷されたデータであるため、読み取りに必要な労力を考慮すると、このような資産運用に資する利用方法はこれまで現実的ではなかった。近年の画像処理技術の進展により、定型・非定型の書類の自動読取が可能になったため、活用が模索され始めたオルタナティブ・データと言える。

上記の例からも示唆されるように、オルタナティブ・データを発掘して活用するためには、運用対象の企業・業界についての深い知識（ドメイン知識）と、大量の非構造化データから情報を効率よく抽出する技術の両輪が必要となる。他社よりも早く未知のオルタナティブ・データを発掘するためには、データ購入費用のみならず、データ分析人材および IT インフラへの投資が欠かせない。推計によると、グローバルでのオルタナティブ・データ関連の支出額は、データ購入、データ分析人材、IT インフラを合わせ、2017 年に 43 億ドルであったのが 2020 年には 90 億ドルと、4 年間で 2 倍に増加する（図表 2）。さらに、先進的な投資技術を競争力とする欧米のクオンツ・ヘッジファンドに加え、特に 2010 年代中盤以降、伝統的な資産運用企業や証券会社ならびに一部のプライベート・エクイティ・ファンド等もオルタナティブ・データの積極的な活用に舵を切ったと報じられている⁶。

³ 一般に、テキスト・データに含まれる対象（企業、経済指標、景気等）について、文章の内容が肯定的（ポジティブ）か否定的（ネガティブ）か、その度合いを数値で表したものを。

⁴ 全体像については佐藤広大「人工知能・ビッグデータを活用した資産運用への期待と課題」『野村資本市場クォーターリー』2017 年春号を参照。

⁵ Greenwich Associates, “Demystifying Alternative Data,” 2019Q2.

⁶ 例として、下記の文献および文献中の参考文献を参照。

饗場行洋、山本裕樹「データサイエンスと新しい金融工学」『財界観測』2018 年春号。

西家宏典、津田博史「従業員ロコミを用いた企業の組織文化と業績パフォーマンスの関係」『証券アナリストジャーナル』2018 年 7 月号。

図表 2 バイサイドのオルタナティブ・データ関連支出（グローバル）



(注) 図表 a のその他費用は人件費、IT 費用等を含む。図表 b は一社当たりのオルタナティブ・データ購入費用。
 (出所) 日本経済新聞⁷、AlternativeData.org、他各種資料より野村証券金融工学研究センター、野村総合研究所作成

2. オルタナティブ・データ注目の背景

伝統的な金融データ以外の情報を用いて市場予測をする試みは、決して最近始まったことではない。古くは古代バビロン王国の商人がユーフラテス川の水量をもとに小麦の生育量を推定し、小麦の市場取引に利用したとされる⁸。現代では、圃場の空撮画像や各種地上センサーからのデータを統合し、農作物の収穫時期や収穫量を精確に推定する試みがなされている⁹。

本来は金融業界での利用を目的に集められたわけではないデータを資産運用に活用しようとする試みに注目が高まっている背景として、第一に、データ分析にかかる計算機技術の飛躍的な進歩ならびに利用可能なデータの劇的な増加、第二に、Mifid II（欧州第 2 次金融商品市場指令）を始めとするリサーチに関する規制の変化、そして第三に、金融危機前後から始まった運用環境の変容が挙げられる。

まず、計算技術の進歩については、CPU やストレージが劇的に低価格化したことが大きい。ブルッキングス研究所のハミルトン・プロジェクトによれば、iPad 1 台相当の計算能力当たりのコンピュータ価格は対数関数的に下がり続けており、2000 年から 2010 年の期間でもおよそ 100 分の 1 に低下している（図表 3）。仮に 1982 年時点の典型的な労働者が現代の iPad に相当する計算能力を持つ何かを購入しようとしたら、賃金 360 年分以上を費やす必要があったとも示されている。加えて、ハドゥープ（Hadoop）¹⁰やスパーク（Spark）¹¹といった分散コンピューティングに特化したオープンソース・ライブラリの登場により、汎用的なハードウェアで大規模なデータ処理を行うことが比較的容易になっている¹²。さ

⁷ 「投資に生きる独自データ ホテル予約・特許…異変読む」日本経済新聞電子版、2019 年 1 月 16 日。

⁸ Andrew W. Lo, Jasmina Hasanhodzic, “The Evolution of Technical Analysis: Financial Prediction from Babylonian Tablets to Bloomberg Terminals,” *Bloomberg Press*, 2010.

⁹ 一例として下記を参照：<https://www.jacom.or.jp/column/2018/03/180327-34917.php>

¹⁰ <https://hadoop.apache.org/>

¹¹ <https://spark.apache.org/>

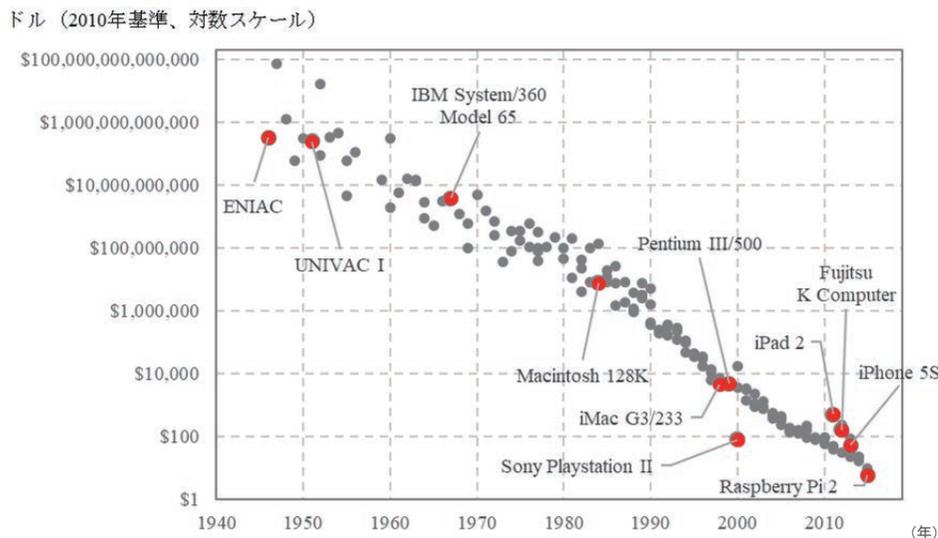
¹² 石井正太「AI とオルタナティブデータ：資産運用の新潮流」MPT フォーラム、2019 年 3 月 7 日。

らに、機械学習技術の進歩、特に深層学習（Deep Learning）の登場により、画像やテキストといった、以前は高精度で分類・構造化を行うことも、意味を解釈することも難しかったデータ群から、何らかの示唆を持つ情報を自動的に抽出することが可能となった。

上述のような技術的な進歩は、1990年代後半以降にインターネットを介したビジネスが普及し、大量のデジタル・データが蓄積されたことで加速している。グーグルやアマゾンなど、デジタル・プラットフォーマーと呼ばれる新興IT企業は、事業から得られたデータをこぞって分析し、自社サービスの改善や広告媒体としての価値向上を図ってきた¹³。しかも、機械学習や大量データの分散処理技術の多くが、オープンソース・ソフトウェアやクラウド・サービスを通じて広く一般に利用可能となってきた。

次に、規制環境の変化としては、Mifid IIによるブローカー・リサーチのアンバンドリングの影響が大きい¹⁴。多くの運用会社にとって、リサーチ・アンバンドリングは外部から提供される投資判断材料の減少につながっており、自社内のリサーチ機能の強化とオルタナティブ・データの利活用を促したと考えられる。運用会社へのアンケートを基にした調査によると、バイサイドのリサーチ収集方法は、セルサイドへの依存が大きい現状から、自社リサーチ、独立系調査会社およびオルタナティブ・データの活用へと移行が大きく進む傾向が報告されている¹⁵。

図表3 iPad 1 台の計算能力あたりの価格



(注) iPad2（2011年販売開始）の1台分の計算能力は1600MIPS（million instructions per second）相当であり、各データポイントは各計算機リリース時点の1600MIPSあたりの価格を示している。

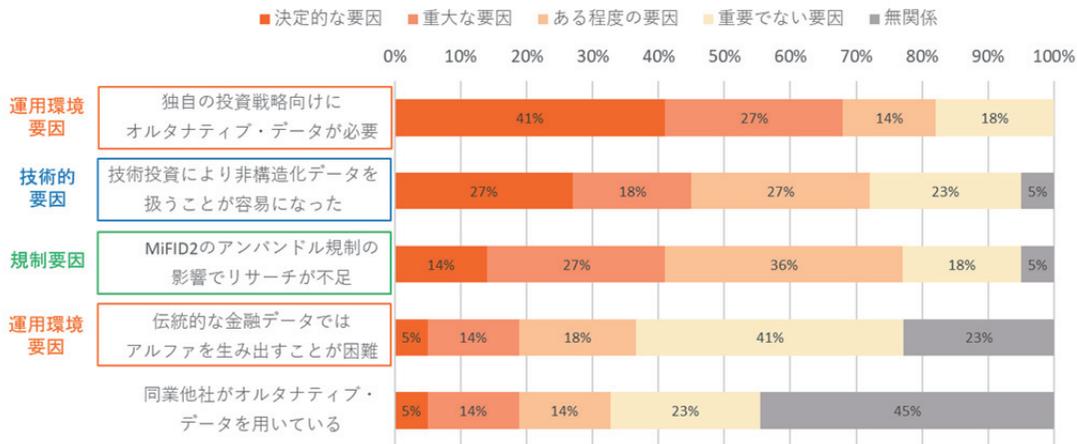
(出所) ブルッキングス研究所ハミルトン・プロジェクト資料から野村証券金融工学研究センター、野村総合研究所作成
元データはモラベック (<https://frc.ri.cmu.edu/~hpm/book97/ch3/processor.list.txt>)

¹³ デジタル・プラットフォーマーと呼ばれる巨大IT企業と、ビッグデータ分析の関連については、伊藤健、佐藤広大「デジタル・プラットフォーマーの戦略とデータの価値を巡る議論ービジネスモデルの特徴と金融サービス業への示唆ー」『野村資本市場クォーターリー』2019年冬号を参照。

¹⁴ Mifid IIの詳細は、神山哲也「欧州委員会によるコミッションのアンバンドリング規制」『野村資本市場クォーターリー』2016年春号を参照。

¹⁵ Greenwich Associates & Thomson Reuters, “Seismic Shifts: The Future of Investment Research,” 2018Q3.

図表 4 オルタナティブ・データ利用の推進要因



(注) アイテ・グループがバイサイド企業に対して行った調査。回答企業数は 22 社。
 (出所) アイテ・グループの調査（公開抄訳）¹⁶から野村証券金融工学研究センター、野村総合研究所作成

運用環境の観点からは、2007 年に発生したいわゆるクオンツ・ショックの際、多くのクオンツ系ヘッジファンドが類似の運用モデルを用いていたために損失が膨らんだとの反省から、独創的な戦略への志向が強まったことが指摘できる¹⁷。そもそも、財務諸表等の伝統的金融データと典型的なクオンツモデルを用いた運用が広く普及したために、独自のデータやアイデアを用いない限り、ファンドマネージャーがアルファを生み出すことが難しくなっていたこともオルタナティブ・データへの関心を高めるきっかけと言えよう。

図表 4 は、バイサイド企業に対してオルタナティブ・データ活用を推進する理由を尋ねた調査の結果である。回答企業の大多数が、技術、規制および運用環境の変化を理由に挙げている。今後、技術の進歩、新しいデータの生成、アンバンドリング規制への世界的な対応などが進むと、オルタナティブ・データの重要性はますます高まると考えられる。

II オルタナティブ・データ導入の効用と課題

1. オルタナティブ・データ利用の目的

オルタナティブ・データの活用には、一般的にウェブやセンサー等から得られるローデータ（生データ）を加工し、情報処理技術を適用できる自然言語（単語や文章）、画像、数値の状態で分析し、そこから運用判断に資する情報を抽出するというプロセスを経る。その際、オルタナティブ・データの有用性を特徴づける要素は主に 3 点挙げられる（図表 5）。

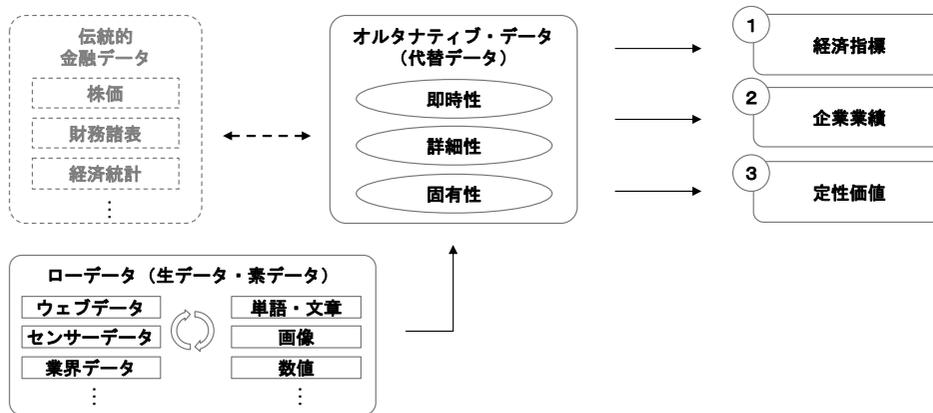
¹⁶ Aite Group, “Alternative Data in Active Asset Management: A New Source of Alpha?” August 14, 2018.

¹⁷ クオンツ・ショックやクオンツ・クエイクと呼ばれる事象については、以下を参照。

Amir E. Khandani, Andrew W. Lo, “What Happened To The Quants In August 2007?: Evidence from Factors and Transactions Data,” *Journal of Financial Markets*, Volume 14(1), February 2011.

「『クオンツショック』から 10 年、AI 進化は本物か」日本経済新聞電子版、2017 年 8 月 9 日。

図表5 資産運用におけるオルタナティブ・データの適用パターン



(出所) 野村証券金融工学研究センター、野村総合研究所作成

1) 即時性

運用判断の材料となる伝統的データとしての経済指標や財務諸表は、集計・監査等の手続きが必要となるため、確定値の公表までには数日から数ヵ月程度のタイムラグが発生してしまうことが多い。また、伝統的データは月次や四半期毎の公表が多く、連続性が低い。この時、オルタナティブ・データを上手に活用することで限りなく実態をリアルタイムに把握できる可能性がある。つまり、多くの市場参加者がまだ認識していないトレンド変化の可視化や、コンセンサス予測値と実態が乖離することで発生するサプライズの予見が期待されよう。

一例として、ブラックロックは、消費者による自動車、航空券、高級品などのウェブ検索データは将来の消費動向を予測するのに有用であると考え、欧州各国の経済見通しを比較分析する際に活用している¹⁸。ただし、この手法は内需予測には有用だが、欧州株式のベンチマーク指標では輸出企業の割合も大きいいため注意が必要であることも記しており、実際の運用においてはオルタナティブ・データを判断材料としつつも、人間による解釈も引き続き重要であることが窺える。

2) 詳細性

既存の経済指標は、何らかの集計値である特性上、対象範囲が国、地域など、一定の単位を持つ。一方、資産運用で投資対象となる株式や債券等の銘柄のエクスポージャーは、必ずしもそれらの指標が映す範囲と一致するわけではない。そこで、メッシュ(網目)化されたオルタナティブ・データにより、詳細な単位で実態を映し出せる可能性がある¹⁹。同様に、既存の財務諸表は、必ずしも分析対象企業で注目されている事業部門毎の実態を詳細に示しているとは限らないので、オルタナティブ・データを用いることで、より詳細に把握できることも期待される。

¹⁸ <https://www.blackrock.com/corporate/literature/whitepaper/bii-interpreting-innovation-us-version.pdf>

¹⁹ 例えば、日本の官民の取り組みとして、地域経済分析システム RESAS (<https://resas.go.jp/>) が挙げられる。

例として、アップルの新型 iPhone などの中核商品の販売動向について考えると、同社の四半期決算では財務諸表に加えて、iPhone、iPad、Mac など商品毎の販売台数と収益が発表されており²⁰、これらの実績値と市場予測との比較が投資家に注目されてきた²¹。ところが、アップルは商品の高級化やサービス収益の成長に目を向けてほしいとの考えから、2018年9月期を最後に商品毎の販売台数と収益は公表しないと²²。クリスマス商戦や新型モデルの販売動向に引き続き注目する投資家は、大手小売店での販売動向を分析するデータ・ベンダーの情報で補完していく試みを続けている²³。このような手法であれば、アップルが元々公表してこなかった地域ごとの iPhone 販売実績などを把握できる可能性もあろう。同じ iPhone でも、容量やサイズ等、複数のモデルが提供される場合は、個別モデルの販売動向分析も有益かもしれない。さらに、アップルのようなメーカーの業績はサプライヤーの業績にも影響することも指摘されており²⁴、その観点でもデータの詳細分析にはニーズがあろう。

3) 固有性

個別の企業が日々行うビジネスは、それぞれの業界固有の情報に溢れている。前述の船荷・貨物情報に類似するデータとして、船舶固有の信号を分析することで物流やコモディティ需給動向を把握する試みも見られる²⁵。

航空業界に目を向けると、航空機が発生する電波を受信することで、地球上を飛んでいる民間飛行機の現在位置をほぼリアルタイムで誰でも把握できるようになっている²⁶。このようなデータを解析することで、物流の把握に留まらず、企業が保有するプライベート・ジェットの航路を予測し、そこから何らかの示唆を得られる可能性も指摘されている²⁷。また、コマツの KOMTRAX²⁸は建設機器に設置されたセンサーをベースにした IoT (モノのインターネット) システムだが、そこから得られた稼働情報の集計値は、月次で地域ごとに公開されている²⁹。このような IoT データは今後さらに増えていく可能性が高いが、様々なセンサーから発信された業界固有データを集計すると、本来の意図に留まらず、使い方次第では資産運用に資する情報源となる可能性がある³⁰。

²⁰ <https://www.businesswire.com/news/home/20181101006102/en/>

²¹ “Apple tumbles 7% after earnings on light guidance,” *CNBC*, November 1, 2018.

²² “Apple will no longer report iPhone numbers after growth went to 0%, and analysts are now worried iPhone sales may decline,” *Business Insider*, November 2, 2018.

²³ <https://media.thinknum.com/articles/the-iphone-xs-xs-max-and-xr-continue-to-lag-at-best-buy/>

²⁴ “Want to Predict Apple’s Stock Price? Watch Its Suppliers.” *Barron’s*, May 3, 2019.

²⁵ “CargoMetrics Cracks the Code on Shipping Data,” *Institutional Investor*, February 4, 2016.

²⁶ <https://www.flightradar24.com/>

²⁷ “Hedge Funds Are Tracking Private Jets to Find the Next Megadeal,” *Bloomberg*, July 2, 2019.

²⁸ <https://sanki.komatsu/komtrax/>

²⁹ <https://home.komatsu.jp/ir/demand-orders/>

³⁰ “Bulldozers Can Show You Where the Economy’s Going Before the Official Data Do,” *Bloomberg*, October 5, 2017.

2. オルタナティブ・データの適用パターン

上述の3つの特徴は、ビッグデータ分析の付加価値とも重複するものだが、オルタナティブ・データ分析の場合、大規模なデータが併せ持つノイズを取り払い、運用パフォーマンスに直結する活用法が目指されていると言えよう。さらに具体的なオルタナティブ・データの適用シーンには、以下が挙げられる。

1) 経済指標の補完

景気や物価の情勢を示す代表的なマクロ指標としては、GDP、雇用統計、消費者物価指数など多数存在する。これらの指標は、投資家の注目度や対象範囲は異なるものの、実体経済を何らかの形で映し出す値である。しかし、前述の通り、公表された値は必ずしも今この瞬間の状態を表すものとは限らない上に、頻度・業種など、どこまで細分化（或いは統合）できるかもまちまちである。ここで、例えばインターネット誕生以来ウェブ上に生成・蓄積されてきた膨大なテキスト情報（文字列）を上手に分析することで、既存指標の即時性を向上・精緻化できるかもしれない。

一例として、英国の国家統計局（Office for National Statistics）では、2014年以來、消費者物価指数を算出するためのデータ収集手段として、オンライン・ショッピングにおける商品価格動向を、プログラムにより、ウェブサイトのソースコードであるHTML（HyperText Markup Language）から直接データ抽出（ウェブ・スクレイピング）する手法等の研究を継続し、指数の試験算出も実施している³¹。そして、2023年には、POS（point-of-sale）データも含めたオルタナティブ・データを実際の消費者物価指数にも織り込んでいくことを計画している³²。別の例として、そもそも適切な統計情報が不足あるいは存在しない領域の経済状況を測定するために、代替的指標を活用することも以前から研究されている。例えばブラウン大学のデイヴィッド・ワイル経済学教授らが2009年に提唱した手法では、宇宙からの撮影により測定した夜間光量を、GDPあるいは電力消費量の補完・代替情報として活用することを試みている³³。このような新しい手法により、既存指標の改良のみならず、細分化や、先行指標の生成に繋がれば、運用判断の一助となろう。

2) 企業業績情報の補完

財務諸表・決算報告書に示される経営状況を、より精緻に、より迅速にとらえる試みは注目される。例えば、水族館を運営するシーワールド・エンターテインメント（SeaWorld Entertainment Inc.）の業績を知る目的であれば、大量のクレジットカード取引データや衛星画像データを分析し、詳細な区分ごと（チケット、商品、食品、駐

³¹ Office for National Statistics, “Research indices using web scraped price data: August 2017 update,” September 1, 2017.

³² Office for National Statistics, “Introducing alternative data sources into consumer price statistics,” May 9, 2019.

³³ J. Vernon Henderson, Adam Storeygard, David N. Weil, “Measuring Economic Growth from Outer Space,” July 2009, Revised in December 2011.

車料等)の売上高を知ることができ³⁴、これは毎四半期の決算報告書に記される情報を巡る解釈を補完するものとなる。また、消費者向けビジネスを行う企業の主力商品・サービスの個別売上動向を探るために、各種ウェブサイトやソーシャル・メディア、モバイル広告、他業界の関連データ等を分析し、業績情報の補完材料とすることも考えられよう³⁵。オンライン・ビジネスの状況に関しては、アプリ関連のトランザクションなど、様々なデジタル・データを通して把握できる。一方、実店舗などのオフライン・ビジネスについては、IoT や各種デバイスの位置情報等の集積により、新規サービス開始に伴う動向変化の把握や、全体傾向を推計することもできよう。

3) 非財務情報あるいは定性的企業価値の可視化

非財務情報あるいは定性的価値の可視化、すなわち決算報告書に必ずしも数値としては開示されない情報の定量化においても、オルタナティブ・データは有用である。

例えば、企業の評判（レピュテーション）やブランド力の測定は主観的になりがちだが、在職者・退職者による口コミ情報が集まるウェブサイトを集積された情報等をベースに自然言語処理を用いてある程度は定量化できる可能性もある。実際、米国の企業評判口コミサイトのグラスドア（Glassdoor）は、投資判断の一助としてヘッジファンドやプライベート・エクイティ・ファームが活用するに留まらず、企業経営者もサイト上の自社従業員の声に注意を払っていると謳われている³⁶。他方、オルタナティブ・データを用いることで企業の ESG（環境・社会・企業統治）や SDGs（持続可能な開発目標）の取り組みを評価し、運用判断に役立てることも期待されており³⁷、クオンツ運用への ESG ファクター取り込みを巡る議論と共に注目される³⁸。

最終的に運用会社が目指すのは、オルタナティブ・データの適用を通じて、伝統的データに基づくファンダメンタル分析を補強することだろう。典型的なファクターモデルがコモディティ化し、インデックス投資などのパッシブ運用も拡大する中、革新的な技術を活用することで、アクティブ運用における投資リターンの向上、特にアルファの導出を実現できるかの挑戦とも言えよう。

3. 多種多様なオルタナティブ・データと有用性

資産運用への活用が期待されるオルタナティブ・データの種類の種類は、技術革新と並行して増加している。例えば、画像データから得られるオルタナティブ・データは、2010 年代

³⁴ 前掲脚注 12 を参照。

³⁵ 前掲脚注 12 を参照。

³⁶ <https://www.glassdoor.com/employers/blog/how-investors-use-glassdoor/>

³⁷ <https://corpgov.law.harvard.edu/2018/11/10/decoding-quant-esg/>

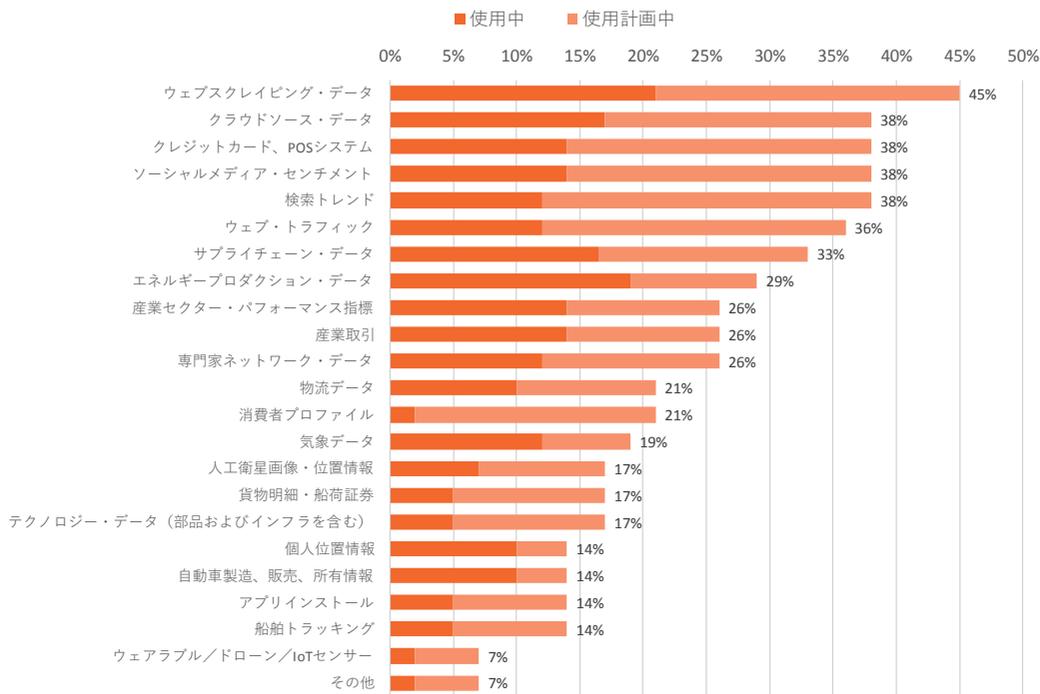
響場行洋、伊藤健、井辺洋平「ESG 格付のネットワーク構造が示す新しい企業戦略 —進化する、定性データの定量化技術—」『証券アナリストジャーナル』2019 年 5 月号。

³⁸ The Principles for Responsible Investment, “A Practical Guide to ESG Integration for Equity Investing,” September 2016.

に起こったディープラーニングの急速な進展によって生成可能となった。具体的事例として、小売店舗の駐車場に停まっている乗用車の台数を衛星画像から画像処理 AI を用いて測定したデータは、小売業者の四半期業績予測に資するオルタナティブ・データとして知られているが、すでに投資情報端末にて配信されており、コモディティ化したとすら言える³⁹。また、携帯電話の位置情報から人の移動に関するデータ（人流データ）を作成し、例えば一定期間にテーマパークを訪れた人数を割り出すといった手法も、従来なかった方法である。他にも多種多様なデータが存在し、AlternativeData.org によれば、現在オルタナティブ・データ・ベンダーの数はグローバルで 400 社を超える⁴⁰。

代表的なオルタナティブ・データを種別に分類し、バイサイド企業へのサーベイ結果を基にそれぞれの使用状況を示したのが図表 6 である。当図表からは、Web サイトやソーシャル・メディアからのデータ、およびクレジットカード・データおよび POS データの人気の高い一方、人工衛星画像や位置情報の順位は低いことが見て取れる。AlternativeData.org の調査においても、最も正確かつ深い知見を得られるデータとして、クレジット／デビットカード・データ、ウェブスクレイピング・データ、ウェブトラフィック・データが上位を占める。反対に、不正確で知見を得にくいデータとして、位置情報および衛星画像データが挙げられている。

図表 6 オルタナティブ・データの使用状況



(注) グリニッチ社がバイサイド企業に対して行った調査より。回答企業数は 42 社。
 (出所) グリニッチ社の調査⁴¹から野村証券金融工学研究センター、野村総合研究所作成

³⁹ “Orbital Insight Opens Its Satellite Network to the Masses,” *Bloomberg*, May 15, 2019.

⁴⁰ <https://alternativedata.org/stats/>

⁴¹ 前掲脚注 5 を参照。

上述のサーベイ結果は、各種オルタナティブ・データの話題性と有用性は必ずしも一致しないことを示唆している。データの有用性を決める要素として、第一に、企業の業績がある程度直接的に推定できるか、第二に、十分な過去データが存在するか、第三に対象企業の多さが挙げられる。例えば、クレジットカード・データや POS データ等の購買データからは、消費者向け商品メーカー等の売上高を相当程度、直接的に推定可能と見做される。一方で、人流データ等から企業業績を推定するためには、データと企業業績との関係に様々な仮定を置いて推定モデルを構築する必要がある。例えば、ショッピングモール内の人出から特定の店舗の売上を予測するには、何人が当該店舗に訪れ、平均いくら消費するかをパラメータとして仮定しなければならない。パラメータの値は曜日、時間帯、季節、天候等に応じて変わることが容易に想像でき、モデルの推定誤差を拡大する要因となる。さらに、人流データや IoT センサーデータ等は比較的最近普及したデバイスから得られるものであるため、バックテストを行うための過去データが高々数年間分しか入手できないことも、導入を困難にする理由であろう。

留意すべき点として、図表 6 において使用率の低いデータは一概に精度や有用性が低いというわけではない。例えば貨物明細・船荷証券などは、POS データではとらえることのできない、B2B の取引量を反映していると考えられ、事業領域によっては唯一無二の情報を得られる可能性がある。重要なのは、検討している投資対象の企業・業種の動向をとらえるオルタナティブ・データを探し出すことと、必要に応じて幅広いデータを相互補完的に用いることと考えられる。

4. 効率的な利用へ向けた課題

オルタナティブ・データ利用に関する課題として、高額なデータ購入費用が挙げられることが多い⁴²。実際、データセット当たり数百万円～数億円という価格は、多くの資産運用企業にとって導入への障壁になると考えられる。一方で、図表 2 に示した通り、オルタナティブ・データの活用を進めるには、いわゆるデータ・サイエンティストやデータ・エンジニアと呼ばれる人材や、大量データを効率よく処理する分析インフラストラクチャーの構築に大きな投資が必要である⁴³。また、新しいオルタナティブ・データの利用を始めるためには、数週間から数か月の期間をかけ、既存の運用戦略に付加価値を与えるかどうかの検証を行うことが必要とされる⁴⁴。

人材獲得については、費用もさることながら、そもそも適切なスキルセットを備えた人材を確保することが難しいのが現実である。欧米では統計学などを専攻した大学卒の初任

⁴² 「投資に生きる独自データ ホテル予約・特許…異変読む」日本経済電子版、2019年1月16日。
「位置情報や SNS 投稿、新顔データで投資先読み」日本経済新聞電子版、2019年5月4日。
「オルタナティブ（代替）データとは 新たな投資分析」日本経済新聞電子版、2019年5月4日。

⁴³ “Investec: For Alternative Data, Learn How to Fail Fast,” *Waters Technology*, July 8, 2019.

⁴⁴ “Alt-Data Difficulties Challenge Largest Asset Managers,” *Waters Technology*, November 7, 2018.

“Buy Side, FISD Begin Industry Push for Alt Data Standards,” *Waters Technology*, May 23, 2019.

給が年間 12 万～20 万ドル、3～5 年の就業経験が加われば 20 万～50 万ドル、データ部門の責任者ならば 100 万～200 万ドルの年収になるとも言われる中、オルタナティブ・データ戦略を考える上では、自社で専門人材の確保が可能かどうかという視点が欠かせない⁴⁵。

オルタナティブ・データに関わるもう一つの課題として、データの適法性に関するリスク管理が挙げられる。個人情報の扱いについて世界的に規制強化の機運が高まるなか、データ利用者としては、購入したデータに個人を特定できる情報（Personally Identifiable Information; PII）が含まれていたり、そもそもデータが不当に収集されたものでないか、懸念を完全に排除したいと考えるであろう⁴⁶。万が一にでもデータに法的な問題があった場合の責任について、事前にベンダー側と詳細に交渉する必要があり、結果としてデータ購入に係る契約合意プロセスが長期化する傾向が指摘されている。

III オルタナティブ・データ供給サイドの多様化

1. 利用者層の拡大とベンダー・コミュニティへの期待

図表 7 では、データの利用目的で分類した、代表的なオルタナティブ・データ・ベンダーを列挙している。大別して、消費者の購買・行動データを用いて消費者向け（To-C）企業の業績を予測するベンダー、ニュースやソーシャル・メディア・データからリアルタイムで株価等のセンチメントを計測するベンダー、衛星画像や各種センサーからコモディティの備蓄・供給量を推定するベンダーに分けることができ、一部を除いて、各ベンダーはいずれかの分野に特化していることが多い。ベンダーごとに得意とするデータ種別や対象資産が異なるため、多くの場合バイサイドは複数のベンダーからデータを購入する必要が生じ、後述する通り、特に伝統的バイサイドがオルタナティブ・データ活用をする際の課題の一つと考えられる。

10 年ほど前から拡大を続けてきたオルタナティブ・データ市場であるが、2～3 年ほど前までは、一部のクオンツ・ヘッジファンドなどだけをユーザーとした、比較的ニッチな分野と考えられていた⁴⁷。しかし、最近になって、伝統的な資産運用企業がオルタナティブ・データの導入を本格的に検討し始めている。利用者層の拡大にともない、データ・ベンダーに期待される役割も徐々に変化している。例えば、一部の先進的なクオンツ・ヘッジファンドは、大量の非構造化データを分析する技術力がすでにあるため、なるべく詳細な情報まで含んだデータ（生データ）を購入し、他社と差別化された取引シグナルを生成したいと考えている。一方で、大多数の投資家は同様のリソースを持たないため、ある程度集約・構造化されたデータ（エクセルで扱えるデータ）を望ましいと考える傾向にある。多くのクオンツ・ヘッジファンドが、四半期毎の企業業績予測等を通じて比較的短期の取

⁴⁵ 「イノベーター 『オルタナデータ』、投資情報に」日本経済新聞電子版、2018 年 8 月 31 日。

⁴⁶ “Alt Data’s Ethical Day of Reckoning,” *Waters Technology*, October 18, 2018.

⁴⁷ 前掲脚注 5 を参照。

図表 7 利用目的別オルタナティブ・データ・ベンダーの例

To-C 企業 業績予測 / 消費者分析

企業	分析データ
Placemeter	スマートシティ、小売り、広告
Yipit Data	オンライン小売サイト集約
MKT Mediastat	消費者行動データ
7Park Data	消費者行動データ
Second Measure	消費者購買データ（クレカ）
Cardlytics	消費者購買データ（小売り、オンライン）
Investnet Yodlee	個人銀行口座・クレカデータ
Airsage	交通量（自動車・歩行者）モニタ
INRIX	交通量（自動車）モニタ
Foursquare	交通量（歩行者）モニタ
Placed	交通量（歩行者）モニタ
Streetlight Data	交通量（自動車・歩行者）モニタ
Datamir	ソーシャル・メディア
Orbital Insight	衛星画像

リアルタイム金融データ

企業	分析データ
Premise	小売り価格、インフレ
Dataminr	ソーシャル・メディア
Selerity	ソーシャル・メディア、ニュース、IR
Quandl	伝統的金融&オルタナデータ
Sentio	伝統的金融&オルタナデータ、ptn-w/ TR, BATS
Owlin	ニュースセンチメント、ptn-w/ ING Bank
Heckyl	ニュースセンチメント他、ptn-w/ Euronext
Ravenpack	ニュースセンチメント、リスクアラート
Discern	各種オルタナデータ

コモディティ

企業	分析データ
Descartes Labs	衛星画像&農業データ、収穫量推計
Genscape	原油、ガス、農業データ
Orbital Insight	原油、中国景気、グローバル収穫量推計
Rezatec	小麦、コメ、コーヒー、鉱物
Rsmetrics	衛星画像⇒交通量モニタ
Spaceknow	衛星画像⇒中国製造業、アフリカGDP
Tellus Labs	衛星画像⇒天候データ、トウモロコシ、大豆
Ursaspace	衛星画像⇒原油

（出所）各種資料から野村証券金融工学研究センター、野村総合研究所作成

引にオルタナティブ・データを利用しようとするのに対し、中長期保有を前提とした機関投資家は、経済の長期見通しや企業の長期成長力への示唆を得ようとする点も異なると言える。ニーズの多様化に対応して、データ・ベンダーには販売データセットの多様化や、オルタナティブ・データの導入支援も行うなど、ビジネス・モデルの拡大・修正を検討する傾向が見られる。

実際、グリニッチ社の調査によると大多数の投資家がオルタナティブ・データを使うために、ベンダーから何らかの支援が必要と考えている。購入データの詳細に関する説明に加え、技術的な支援やデータ・サイエンス（データ分析）についての支援を求めるバイサイド企業も一定数見受けられると報告されている⁴⁸。なるべく少数のベンダーから、用途に適したデータの提案や必要に応じた技術的な支援を得つつ、同時に法的リスクも制御しながらオルタナティブ・データの導入を進めたいとの意向が示唆される。

2. 大手情報ブローカーの参入とベンダー・マーケット・プレイス

データ購買側のニーズの変化に呼応する形で、販売側の動向も変化している。注目されるのは、ブルームバーグやファクトセット等の大手金融情報ベンダーによる、マーケットプレイス型のオルタナティブ・データ提供サービスであろう。ブルームバーグはエンター

⁴⁸ 前掲脚注 5 を参照。

プライズ・アクセス・ポイント⁴⁹、ファクトセットはオープン・ファクトセット・マーケットプレイス⁵⁰という名称で展開しており、両者とも、数あるオルタナティブ・データから選りすぐったものを、自社の情報端末やウェブサイトを通じて販売・提供する事業である。

購買するバイサイドの視点から見たマーケットプレイス型サービスの利点として、第一に、一定程度質の担保されたデータが一か所に集まっており必要なデータを探しやすいこと、第二に、大手情報ベンダーが提供する伝統的データとオルタナティブ・データを同時に使いやすいこと、第三に、購買契約を大手情報ブローカーに一元化でき法的リスクも管理しやすいことが挙げられるだろう。特に第一と第二の利点については、データの基礎的な検証や伝統データとのマッチングを購入する側が行わなくて済むという点で、オルタナティブ・データ利用の障壁を大きく下げるものと考えられる⁵¹。

3. 取引所や格付け機関によるデータビジネスの拡大

金融情報ベンダー以外では、大手取引所や格付け機関による類似の取り組みが注目されるよう。2018年にはナスダック取引所が大手のオルタナティブ・データ・ベンダーであるクアンドル（Quandl）を買収して関心を集めた⁵²。また、同年には、オルタナティブ・データ領域の強化を進める S&P グローバルが、金融専門の AI ベンチャーであるケンショー・テクノロジーズ（Kensho Technologies）を買収するなど、近年、大手プレーヤーによるデータ・ベンダーやデータ分析企業の買収が加速している（図表 8）。

このような動きが台頭している背景としては以下の点が挙げられる。第一に、収益源の多角化である。取引所や格付け会社はそれぞれの本業である金融商品の取引や発行体の信用格付けを行う中で、大量の情報を取り扱うため、自ずと付随情報を顧客に提供するサービスの拡充に取り組んできた。現在では、データ提供自体がビジネスとして重要な位置づけとなっている。例えば、ナスダックの場合、情報サービス部門は市場データ、インデックス、投資データ・分析の各サービスから構成されているが、2018年の収益全体に占める同部門の割合は 28.3%となっており、IPO 関連収益を含むコーポレート・サービス部門（20.9%）を上回り、取引関連手数料を含む市場サービス部門（37.9%）に次ぐ柱となっている⁵³。同部門収益額の前年比成長率（21.4%増）は全部門で最も高く、戦略上も最大の成長機会を見込む領域と位置付けられている。

⁴⁹ <https://eap.bloomberg.com/>

⁵⁰ <https://open.factset.com/en-us>

⁵¹ “The New Oil Rush: An Examination of the Alternative Data Market,” *Waters Technology*, October 25, 2018.

⁵² “Nasdaq Buys Quandl to ‘Dominate’ Alt Data Space,” *Waters Technology*, December 5, 2018.

⁵³ Nasdaq, Inc. 2018 Form 10-K を参照。

図表 8 代表的なオルタナティブ・データ関連 M&A 事例

発表時期	買い手 (出資元企業)	被買収企業/部門(出資先企業)	金額
2019年8月	LSE	Refinitiv (旧 Thomson Reuters Financial & Risk unit)	270億ドル
2019年7月	Moody's	Four Twenty Seven (気候データ)	非開示 (過半数取得)
2019年2月	S&P Global	Panjiva (貿易データ)	非開示
2018年12月	Vista Equity Partners	7Park Data (非構造化データ)	非開示
2018年11月	Nasdaq	Quandl (金融データ)	非開示
2018年6月	Warburg Pincus	Reorg Research (ディストレスト関連データ)	3.75億~4億ドル
2018年5月	Fitch	Fulcrum Financial Data (ディストレスト関連データ)	非開示
2018年5月	Moody's	QuantCube (金融 AI)	500万ドルラウンドにおけるマイノリティ出資
2018年5月	IHS Markit	Ipreo (マルチアセットデータ・ソフトウェア)	18.55億ドル
2018年3月	S&P Global	Kensho Technologies (金融 AI)	5.5億ドル
2018年1月	Blackstone	Thomson Reuters Financial & Risk unit (財務・リスクデータ)	170億ドル (55%取得)
2017年9月	Nasdaq	eVestment (運用会社向けデータ)	7億ドル
2017年5月	Moody's	Bureau van Dijk (ビジネスデータ)	30億ユーロ
2016年11月	LSE	Mergent (上場・未上場企業データ)	非開示 (100%取得)
2016年10月	Morningstar	PitchBook (未上場市場データ)	1.8億ドル (20%保有済みで残り 80%を買収)
2016年5月	General Atlantic	Argus Media (コモディティデータ)	10億ポンド
2016年2月	IHS	OPIS (コモディティデータ)	6.5億ドル
2015年10月	ICE	IDC (金融データ)	52億ドル
2015年7月	McGraw Hill Financial	SNL Financial (金融データ)	22億ドル
2014年4月	Blackstone, Goldman Sachs Merchant Banking	Ipreo (マルチアセットデータ・ソフトウェア)	9.75億ドル (報道情報)

(出所) 各種公開情報から野村証券金融工学研究センター、野村総合研究所作成

第二に、データの中立性である。取引所や格付け会社は資本市場全体に資する機能を提供するインフラ的な役割を果たしており、元来、個別の金融機関や投資家、発行体等とは公平に接する中立的な存在と言えるが、LIBOR で発生したような不正操作あるいは利益相反の懸念を解消するためにも、金融市場関連データが中立的データ・プロバイダーに集められる動きが加速したと考えられる。

第三に、プライベート・エクイティ (PE) ファームの動きである。図表 8 の通り、データ・ベンダーや企業のデータ部門が分離・統合し、ワンストップのプラットフォームへと発展することで、付加価値が向上し、評価額も上昇する傾向が見て取れるが、大型ディールの背後には PE ファームが存在するケースも多い。例として、金融データ・プラットフォームのアイプリオ (Ipreo) を見ると、同社は PE ファームのヴェロニス・スラー・スティーヴンソン (VSS) が 2006 年 12 月に前身となる 2 社を統合する形で設立されたが⁵⁴、2011 年 5 月に KKR が 4.25 億ドル (報道ベース) で買収したのち⁵⁵、2014 年 4 月にブラックストーンとゴールドマン・サックスのマーチャント・バンキング部門が

⁵⁴ <https://ipreo.com/press-releases/i-deal-and-hemscott-merge-to-form-ipreo-announce-first-acquisition-as-combined-entity/>

⁵⁵ "KKR receives PE bids of close to \$800m for Ipreo," *Financial Times*, February 25, 2014.

9.75 億ドル（報道ベース）で買収した⁵⁶。さらに 2018 年 5 月にアイプリオを 18.55 億ドルで買収⁵⁷した IHS マークイット（IHS Markit）も、自身がデータ・プロバイダー同士の統合を繰り返して大規模なデータ会社へと成長してきたプレーヤーである。同様に、2018 年 1 月にブラックストーンはトムソン・ロイターからスピノフしたファイナンシャル・リスク部門の過半数株式を部門評価額 200 億ドルのもとで取得したが⁵⁸、同部門はリフィニティブへと名称変更されたのち、2019 年 8 月に 270 億ドルでロンドン証券取引所（LSE）グループに売却されることが発表された⁵⁹。また、2015 年 10 月に発表されたインターコンチネンタル取引所（ICE）による 52 億ドルでのインタラクティブ・データ・コーポレーション（IDC）の買収案件は、PE ファームのシルバーレイクとウォーバーク・ピンカスが売り手となった⁶⁰。

4. セルサイド金融機関の動向

データ販売側の立場で、オルタナティブ・データ関連ビジネスを推進する大手金融機関も増えている。ゴールドマン・サックスは、トレンド分析、プライシング分析、クレジット・デフォルト・スワップ（CDS）スプレッドの取引データ等、従来は社内用と考えられてきたデータを顧客に提供することで収益化する構想を示している⁶¹。情報の機密性に対する顧客の懸念に注意を払いつつ、サブスクリプション（定額継続課金）型のデータ・アズ・ア・サービス（data-as-a-service）を模索していると報じられている⁶²。

また、シティグループは、データ・スタートアップのシンクナム（Thinknum）と提携して、決算書や経済指標などの伝統的な情報源からは容易に得られない、企業の経営状況に関連するオルタナティブ・データを、顧客に提供し始めている⁶³。シンクナムが取りまとめるデータとしては、ブランド力を測るソーシャル・メディアのトラフィック情報や、企業の求人件数、オンライン販売される商品の価格及び入手可能性等が挙げられるが、リサーチ情報等を提供するウェブ・プラットフォームのシティ・ベロシティ（Citi Velocity）を通して、そのようなオルタナティブ・データの一部に、顧客であるヘッジファンドや資産運用会社がアクセスできるようにしている。加えて、シティグループでは顧客の個別要望に応じて、徒歩交通、携帯電話シグナル、人工衛星画像、クレジットカード及び POS 取引データ、位置情報、集約化された法人資金フローなど、非市場情報を取

⁵⁶ “Blackstone, Goldman Sachs to buy KKR's Ipreo,” *Reuters*, April 14, 2014.

⁵⁷ <https://news.ihsmarkit.com/press-release/corporate/ihs-markit-completes-acquisition-ipreo>.

⁵⁸ <https://www.thomsonreuters.com/en/press-releases/2018/january/thomson-reuters-and-blackstone-announce-strategic-partnership-for-thomson-reuters-financial-and-risk-business.html>

⁵⁹ <https://www.londonstockexchange.com/exchange/news/market-news/market-news-detail/LSE/14171701.html>

⁶⁰ <https://ir.theice.com/press/press-releases/all-categories/2015/10-26-2015-135133268>

⁶¹ “Goldman building team to sell its own alternative data,” *Risk.net*, May 2, 2018.

⁶² “Goldman Sachs is exploring plans to create a Netflix for data, and it marks a new frontier for Wall Street,” *Business Insider*, May 28, 2019.

⁶³ “Citigroup has found a new way to offer hedge funds obscure data that can give them an edge — and it's part of a \$2 billion investing gold rush,” *Business Insider*, September 15, 2018.

り扱い、データの提供やテストをサポートするコンサルティング体制を整備していることが報じられている⁶⁴。

その他の大手金融機関では、バークレイズ、クレディ・スイス、JP モルガン・チェース、モルガン・スタンレー、UBS も、資産運用会社の顧客向けにオルタナティブ・データ提供を試みていると見られる⁶⁵。主な課題として、提供されるデータの品質、データの所有権並びに秘匿性、一部の顧客に特別な情報を提供していると誤解されることへの懸念が挙げられるが、推進各社は、顧客や業界・市場全体に資するサービス提供を目指して、慎重に取り組んでいる状況と考えられよう。

IV 伝統的な資産運用プレーヤーへの示唆

資産運用の観点では、運用モデルのコモディティ化により他者との差別化が図りにくくなっている中、オルタナティブ・データの活用がアルファの源泉になると期待される。もともと、オルタナティブ・データを用いれば必ず運用パフォーマンスが向上するわけではなく、あくまで土台となるファンダメンタル分析を補強する材料と捉えるべきであろう。また、オルタナティブ・データの活用の際に、他の投資家が市場をどう見ているか、つまり、現在注目されている投資テーマが何なのか、ある情報が市場の価格に織り込み済みなのか、あるイベントに関する実績値とコンセンサス予測値に乖離は生じるのか等を考慮することの重要性は変わらない。そのため、先進的なクオンツ・ヘッジファンドだけでなく、伝統的な資産運用プレーヤーにとってもこれまで蓄積した運用ノウハウを新たな形で活かせる機会となる。

他方、データ取り扱いの巧拙でも差は生まれてくると考えられる。具体的には、上述のように運用判断の中でオルタナティブ・データをどう活用するかだけでなく、ターゲットとなるデータはどうすれば取得できるか、そのためのデータ購入費用や加工・分析コストにはどれほど費やせるか、最終的にデータは使い物になるのかといった技術力や目利き力が試される。何よりも、注目されるオルタナティブ・データは、やがて多くの市場参加者が活用しだし、自ずとコモディティ化していく宿命にあり、結果としてアルファが消失してしまうことも考えられる。そのため、一般に専門性あるいは集中度合いの高い運用スタイルにとって重要となるキャパシティの制約 (capacity constraints) や投資戦略の機密性保持に注意を払いつつ、他者が目をつけていないオルタナティブ・データを探したり、複数のデータセットを組み合わせ活用したり、独創的な適用の仕方を考えたりする競争は続くだろう。そのような競争下では、自社の強みを、適切なパートナーとの協業により強化・補完していくことが選択肢となりうる。オルタナティブ・データの活用は、バイサイド、セルサイド、データ・ベンダー等の関係者が一丸となり取り組むことが期待されよう。

⁶⁴ “Banks discreetly seek personnel to mine alt data riches,” *Risk.net*, September 4, 2018.

⁶⁵ “The New Oil Rush: An Examination of the Alternative Data Market,” *Waters Technology*, October 25, 2018.