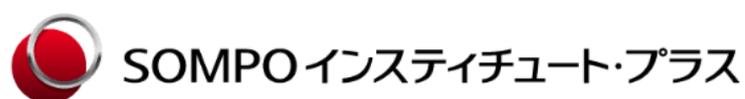


# オルタナティブデータ FACTBOOK

2022年9月



## はじめに

この度、一般社団法人オルタナティブデータ推進協議会（以下、推進協議会）からの委託を受け、弊社 SOMPO インスティテュート・プラスが推進協議会の会員に向けてアンケート調査を実施し、オルタナティブデータの活用状況に関する FACTBOOK を取りまとめることができました。初めに、調査にご協力いただいた会員の皆様、また、集計作業等をサポートいただいた推進協議会の関係者の皆様に対し、厚く御礼を申し上げます。

この FACTBOOK と同様のサーベイの集計結果は、オルタナティブデータの活用で日本に先行する欧米先進国では見かけますが、国内においてこうした包括的な調査が実施され、公表されるのはおそらく初めてのことでしょう。近年、日本においても、伝統的なデータにはないオルタナティブデータ固有の価値が認識され始め、その活用例や活用分野は急速な広がりを見せています。今後そうした流れを定着させ、さらに発展させていくためには、①まずは現状のデータの活用状況を広く共有し、②将来の活用拡大に向けた取り組みや、その過程でデータの提供者や利用者が直面する課題を正しく認識することが不可欠です。そうした意味で、この FACTBOOK を作成、公表することの意義は大きいと考えています。

今後、こうしたアンケート調査の実施を定期的に続けることで、推進協議会の現会員の皆様はもとより、これからオルタナティブデータを活用することを検討中の方々の間で知見・経験が一段と蓄積・共有され、オルタナティブデータの推進に大いに役立つのではないかと考えています。次回以降の質問形式や内容などについて、皆様のお考えやご要望をお聞かせいただければ幸いに存じます。

今回の FACTBOOK から読み取れるとおり、日本のオルタナティブデータ活用は緒に就いたばかりであり、基本的には、購入者、データプロバイダー、リサーチ全ての部門で、その市場規模や用途の幅は広がっていくものと予想されます。その一方で、昨今の外部環境の変化を考えると、国内外ともに、オルタナティブデータの活用の在り方が「転機」を迎える可能性も出てきていると思います。

第一に、新型コロナウイルス感染症の経済・社会への影響が収束に向かうことに伴う、高頻度データに対するニーズの減少です。近年、国内外でオルタナティブデータ活用の流れが急速に高まった背景としては、何と云っても 2020 年からの新型コロナの流行が大きかったと考えられます。感染症の流行は直線的には進まず、変異株の出現を伴いながら、「感染拡大」→「一服・沈静化」→「再拡大」という目まぐるしいステップを繰り返し辿りました。そのことから、地区・時間帯別の滞在人口データやクレジットカード・モバイル決済手段に基づく消費支出データなど、ごく短期間のうちに急激に変化する社会情勢、景気局面をタイムリーに把握できる高頻度データの利用価値が高まりました。これからも新たな変異株や別の病原体が出現する可能性まで考えると、国際経済・社会が感染症の脅威に対して脆弱な面は残っています。しかし、今回の COVID-19 の影響に限れば、中国など極端なゼロ・コロナ政策を採る国を除いては、既に経済・社会的な活動を強く制約する局面は過ぎているように窺われます。このまま感染症の影響がさらに弱まっていくならば、高頻度データに対する需要も一旦は下火になっていくでしょう。

第二に、データの活用に伴って必然的に発生する、個人情報や知的財産権の保護に関する議論や法制的整備が、今後一段と進むと考えられます。これは、公的統計など伝統的なデータの活用においても古くからみられた問題ですが、民間企業がデータ提供主体となるケースが多いオルタナティブデータについては、その利用拡大に伴って、従来以上に個人や個別企業の情報管理、コンプライアンス遵守の観点からの対応が問われることが増えています。そうした点を関係者が強く意識し、今後の課題であると感じている点は、今回の FACTBOOK から明確に読み取ることができます。

もっとも、これらの「転機」は、オルタナティブデータの推進にとって、必ずしも制約的な要素になるとは限りません。むしろ、機を捉えて的確な対応を行っていけば、長い目でデータの活用余地やデータに対する需要の裾野を広げていく「好機」にもなると考えています。

第一の点で挙げた高頻度データの利用価値について言えば、近年ますます増えつつある豪雨や地震といった自然災害の発生時に、企業の生産拠点やサプライチェーンの被害・復旧状況をいち早く把握することができるなど、感染症下の経済でなくても高頻度データの有用性が認められるケースは少なからずあると考えられます。また、私自身が専門とするマクロ経済分析の分野では、ごく足もとの消費支出を追いかける目的でオルタナティブデータを用いるだけでなく、属性別にみた人々の行動をパネル分析するといった形の応用例も増えています。例えば、消費者の年齢やITリテラシーの程度などによって、オンライン消費の定着率がどう異なるか、といった研究です。さらに、毎日の感染状況に関するニュースに馴染んだ一般の方々には、「オルタナティブデータ＝人出情報」と思われがちかもしれませんが、推進協議会の会員の皆様ならばご理解いただけるように、実際のオルタナティブデータは、大量の金融データを含む高粒度データ、ニュースソースや企業決算における様々な開示情報といったテキストデータなど、その定義と範囲は極めて広く、多彩です。その点も今回のFACTBOOKには如実に表れています。現在の局面は、そうしたオルタナティブデータの多様性をアピールし、潜在的なニーズをさらに掘り起こすチャンスの時でもあります。

第二の情報管理の点について言えば、政府は昨年秋にデジタル庁を発足させ、DFFT（Data Free Flow with Trust、信頼性のある自由なデータ流通）を合言葉に、各種データ・情報の社会的な活用に道を拓こうとしています。ここで私たちのような民間企業に必要なことは、政府の対応を受け身で待つのではなく、データの提供者と利用者がともに安心してデータを利活用できる仕組みはどうあるべきか、最近の成功例や苦心談を具体的に示しつつ、政府や関係者に積極的に働きかけていくことだと思います。

今回のFACTBOOKが、そうしたオルタナティブデータ業界の持続的な成長・発展のための一助となることを願っています。

SOMPO インスティテュート・プラス株式会社  
プリンシパル兼エグゼクティブ・エコノミスト 亀田制作

## 1. 本調査とオルタナティブデータ推進協議会について

### (1) オルタナティブデータとは

オルタナティブデータとは、デジタル化の進展などを背景に、金融機関や一般企業、エコノミスト等が新たに利用することが可能になった様々なデータの総称であり、公的統計など伝統的なデータと区別するための用語です。オルタナティブデータの代表的な例として、POS 売り上げデータや経済ニュース、気象情報や位置情報などがあります。このような従来は活用が難しかったデータ群が、機械学習や自然言語処理の発展、コンピュータの性能向上などによって注目を集めるようになっていきました。オルタナティブデータは、もっぱら金融機関や投資家が資産運用を効率化するために参照されてきましたが、近年では事業会社による事業開発や公的機関による社会情勢の把握に活用されるなど、活躍の場が急速に広がっています。

図表 1—1 オルタナティブデータとトラディショナル（伝統的）データ



### (2) オルタナデータ推進協議会について

世界で活用が進んでいるオルタナティブデータですが、日本ではレギュレーション、人材不足、コストといった点で課題が存在し、活用は道半ばとなっています。こうした課題を業界で一丸となって解決すべく立ち上がったのが、「一般社団法人オルタナティブデータ推進協議会（以下、JADAA）」です。現在、金融機関やデータプロバイダー、データ分析企業など多くの参加者が活動事例や課題の共有を行い、オルタナティブデータの利活用促進に取り組んでいます。

### (3) 本調査について

海外ではオルタナティブデータに関する知見が蓄積しつつあり、利用実態についての調査研究も進展しています。一方で、日本では「どのようなオルタナティブデータが」、「どのような業種で」、「どの程度利用されているか」といった実情が現在のところ明らかになっていません。

本調査では、国内のオルタナティブデータの活用状況や課題を明らかにするために JADAA の会員に対し、オンラインでアンケートを実施し、FACTBOOK として取りまとめました。本調査は一般に広く調査した結果ではありませんが、オルタナティブデータを実際に用いている、あるいは利用に強い関心がある企業に聴取を行っており、当業界の現状と課題を的確に示していると考えられます。本調査は今後も継続的に行い、日本のオルタナティブデータ動向を定点観測していく予定です。

#### (4) 本調査の回答者について

本調査は、JADAA の会員にアンケートを依頼し、下記の通り、総計 52 社からの回答を得ました。

図表 1 - 4 本調査の回答者

業種	回答数	割合
情報通信業	26	50%
金融・保険業（資産運用業のぞく）	9	17%
資産運用業	4	8%
学術研究、専門・技術サービス業	7	13%
その他	6	12%
総計	52	100%

#### (5) 本調査の集計期間について

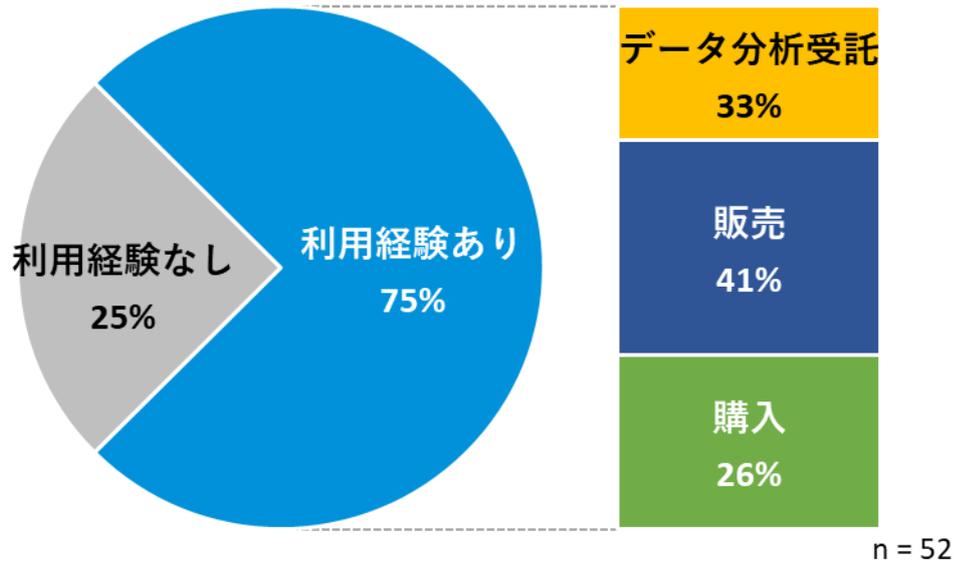
本調査は、2022 年 5 月 30 日～2022 年 6 月 29 日に集計を行いました。

## 2. オルタナティブデータに関する基本情報

### (1) オルタナティブデータの利用経験

オルタナティブデータの利用経験の有無に対する質問に対しては、75%が「利用経験がある」と回答しました。さらに利用経験があるとした回答者は、「購入(26%)」、「販売(41%)」、「データ分析受託(33%)」に分かれています。

図表 2-1 オルタナティブデータの利用経験はありますか？



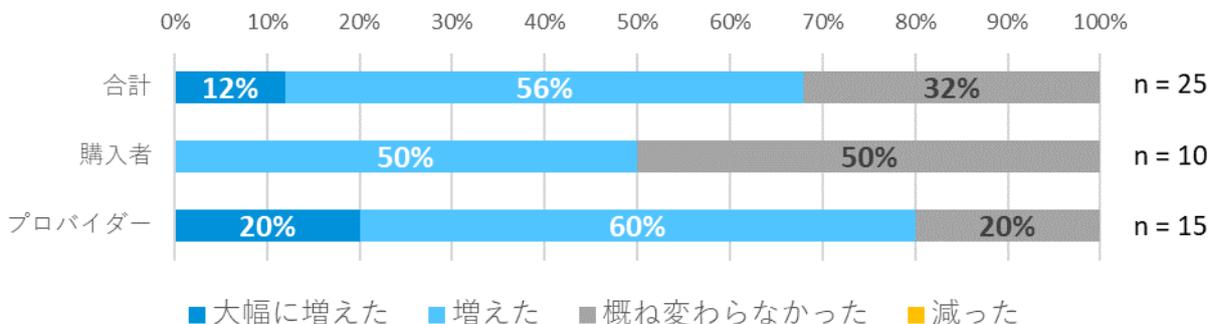
以下では、購入経験がある者を「購入者」、販売経験がある者を「データプロバイダー」、データ受託経験がある者を「データ受託者」、利用経験がない者を「未利用者」と定義します。

### (2) オルタナティブデータの取扱額の昨年度比

オルタナティブデータの年間取扱金額が、前年度に対して「大幅に増えた(12%)」、「増えた(56%)」となり、増加傾向にあることが明らかになりました。

内訳をみると、購入者では「増えた(50%)」、データプロバイダーでは「大幅に増えた(20%)」、「増えた(40%)」となっており、データプロバイダーの引き合いが強いことがわかります。

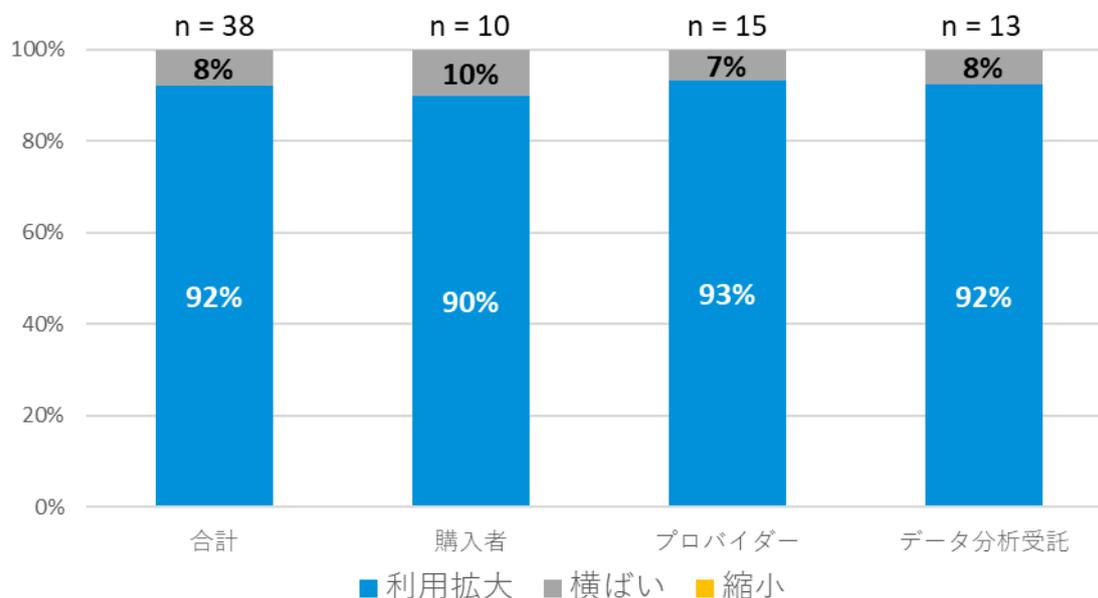
図表 2-2 2022年度のオルタナティブデータの取扱額は、前年度比でどう変化しましたか？



### (3) オルタナティブデータ利用の今後の見通し

今年度含む向こう3年程度のオルタナティブデータ利用方針では、「利用拡大(92%)」が多数を占めました。購入者、データプロバイダー、データ分析受託の主体別でも同様の傾向となりました。

図表2-3 今後(今年度含む向こう3年程度)のオルタナティブデータの利用開始方針は？

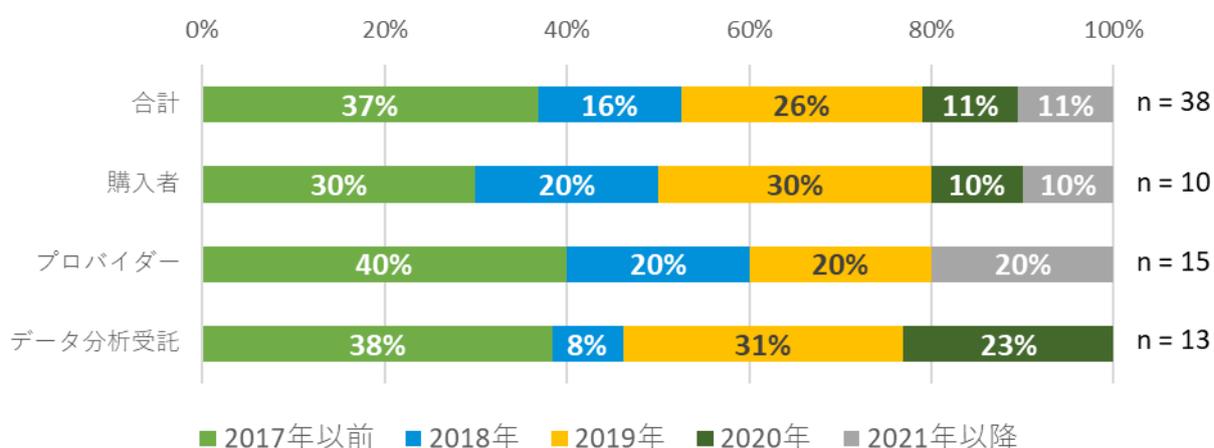


### (4) オルタナティブデータの活用開始時期

オルタナティブデータの活用開始時期は、全体では「2017年以前(37%)」が最多となりましたが、2018年以降も参入が継続しています。

購入者やデータプロバイダー、データ分析受託の内訳をみても、上記の傾向と一致していますが、データプロバイダーの利用開始時期がやや早い傾向にあるようです。

図表2-4 オルタナティブデータの利用開始時期はいつですか？



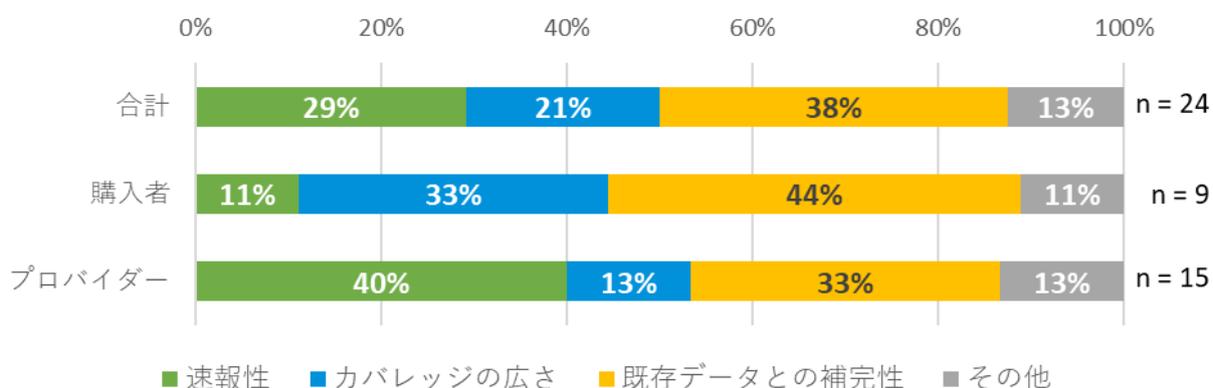
### (5) オルタナティブデータを活用する利点

オルタナティブデータを利用するメリットは、リアルタイムにデータが確認できる速報性や、公的統計に比べたカバレッジの広さなどが挙げられます。全体のアンケート結果では、「既存データとの補完性 (38%)」が最多で、続いて「速報性 (29%)」、「カバレッジの広さ (21%)」の順番となりました。

内訳をみると様相が少し異なります。購入者では「既存データとの補完性 (44%)」、「カバレッジの広さ (33%)」が上位を占めました。データプロバイダーでは「速報性 (40%)」が最多を占めました。

また、その他の回答として購入者では「研究対象として利用 (11%)」、データプロバイダーでは「データの希少性 (11%)」、「既存データとの差別化 (11%)」もありました。

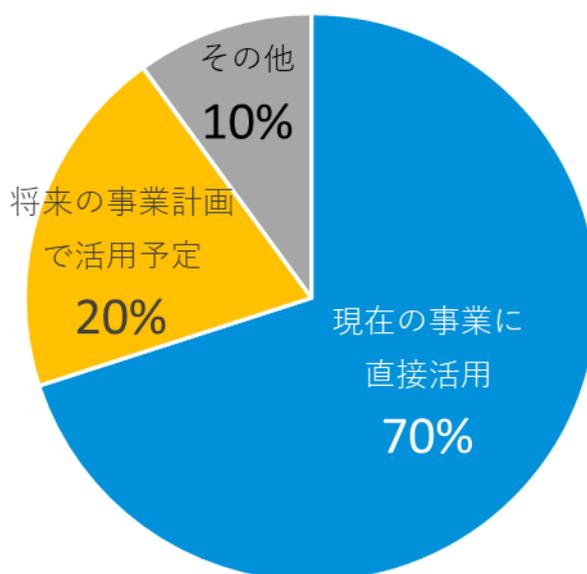
図表 2-5 オルタナティブデータを利用する最大のメリットは何ですか？



### (6) オルタナティブデータの活用目的

購入者を対象に、購入しているオルタナティブデータをどのように活用しているか質問しました。「現在の事業に直接利用している (70%)」、「将来の事業計画で活用する予定がある (20%)」、「その他 (10%)」との回答を得ました。その他の具体的な内容について、「学術研究目的で利用している」との回答でした。

図表 2-6 オルタナティブデータをどのように活用していますか？

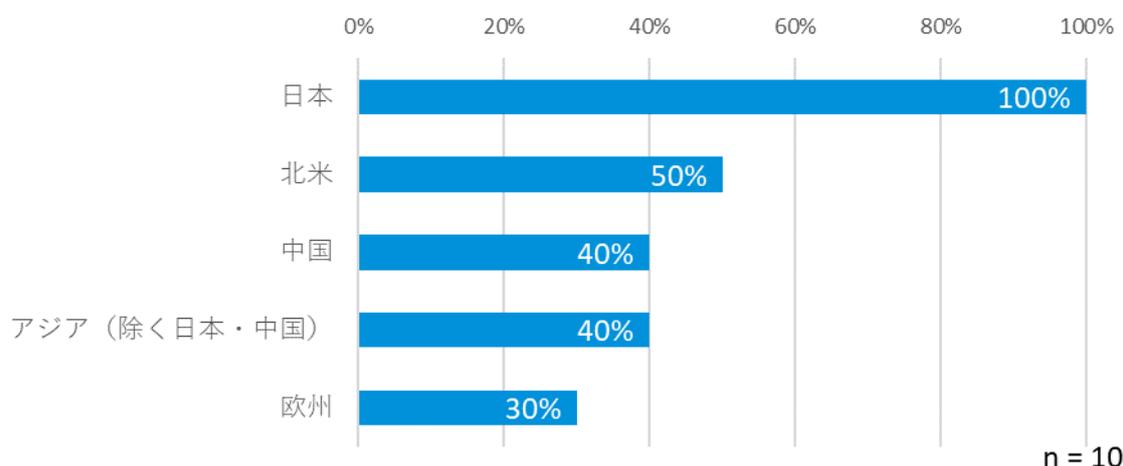


n = 10

### (7) 購入しているオルタナティブデータの対象地域

購入者に対して、購入しているオルタナティブデータの対象地域を確認したところ、日本が最多(100%)となりました。海外では、北米(50%)、中国(40%)、日本と中国を除くアジア(40%)、欧州(30%)となり、その他の選択はありませんでした。海外については、経済規模の大きな地域を中心にオルタナティブデータが購入されていると考えられます。

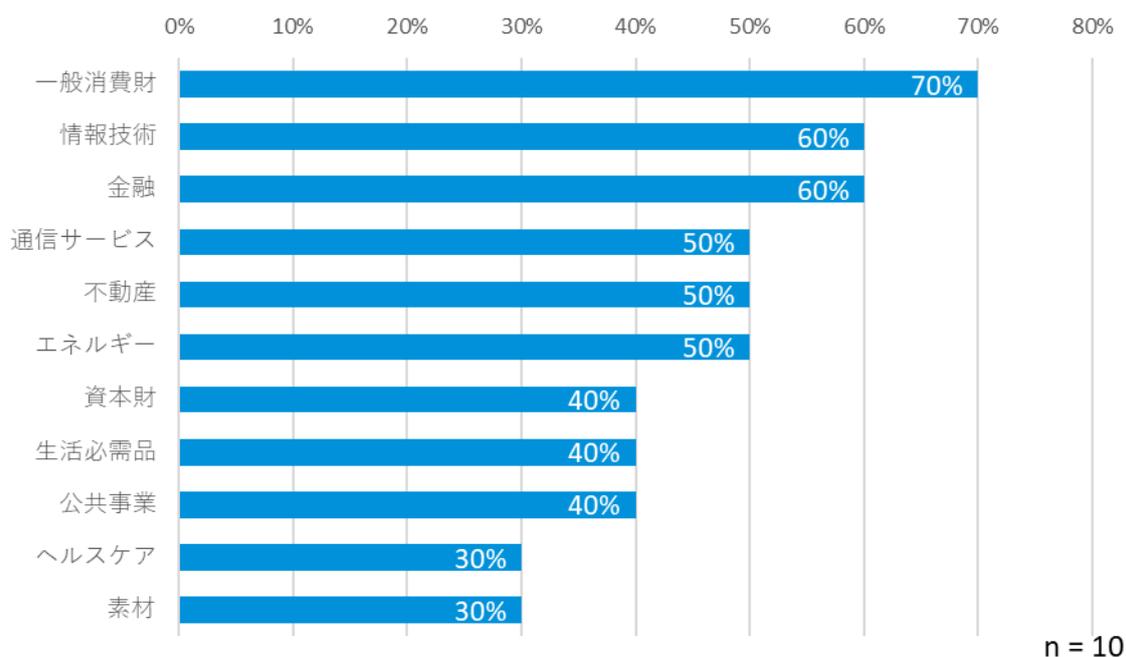
図表 2-7 購入しているオルタナティブデータの地域はどこですか？(複数選択可)



### (8) 購入しているオルタナティブデータの対象セクター

購入者に対し、購入しているオルタナティブデータの対象セクターを集計したところ、一般消費財(70%)、情報技術(60%)、金融(60%)となりました。

図表 2-8 購入しているオルタナティブデータの対象セクターは何ですか？(複数選択可)



### (9) 取扱っているオルタナティブデータの種類

オルタナティブデータは、POS データやクレジットカードデータ、位置情報データをはじめとして、無数の種類が存在します。取扱っているデータの種類のについて、購入者、データプロバイダー、データ受託者にそれぞれ質問しました。

全体では、「位置情報データ (50%)」が最多となり、「POS データ (36%)」、「ニュースデータ (36%)」、「SNS データ (33%)」が続きました。

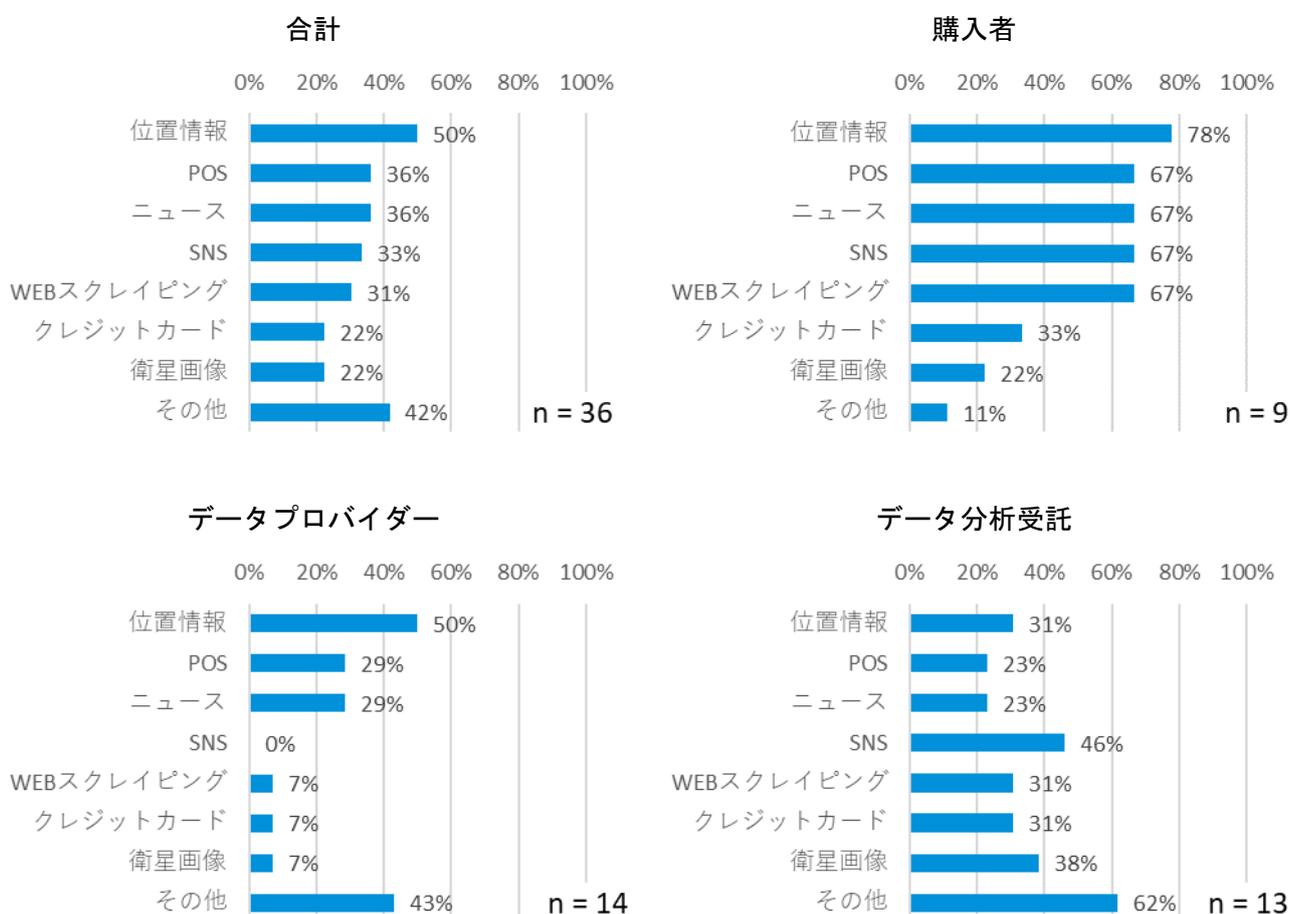
一方で、内訳をみるとそれぞれで取扱うデータの種類の大きく異なります。購入者では、「位置情報データ (78%)」、「POS データ (67%)」、「ニュースデータ (67%)」、「SNS データ (67%)」、「WEB スクレイピングデータ (67%)」が多く利用されています。

データプロバイダーでは、「位置情報データ (50%)」、「POS データ (29%)」、「ニュースデータ (29%)」と購入者に比較してばらつきが大きくなりました。

データ分析受託では、「POS データ (46%)」、「衛星画像データ (38%)」、「位置情報データ (31%)」、「WEB スクレイピングデータ (31%)」、「クレジットカードデータ (31%)」の順番となりました。

また、その他のデータとして、「サプライチェーンの取引関係 (購入者)」、「レセプトデータ、健康診断データ (データプロバイダー)」、「社員ロコミデータ (データプロバイダー)」、「金融取引データ (データ分析受託)」、「気象データ (データ分析受託)」、「WEB 閲覧データ (データ分析受託)」、「共通ポイントカードデータ (データ分析受託)」などの多様なオルタナティブデータを取扱っていることがわかりました。

図表 2-9 利用しているオルタナティブデータの種類の？ (複数選択可)



### (10) 取扱開始時期と扱っているデータの種類

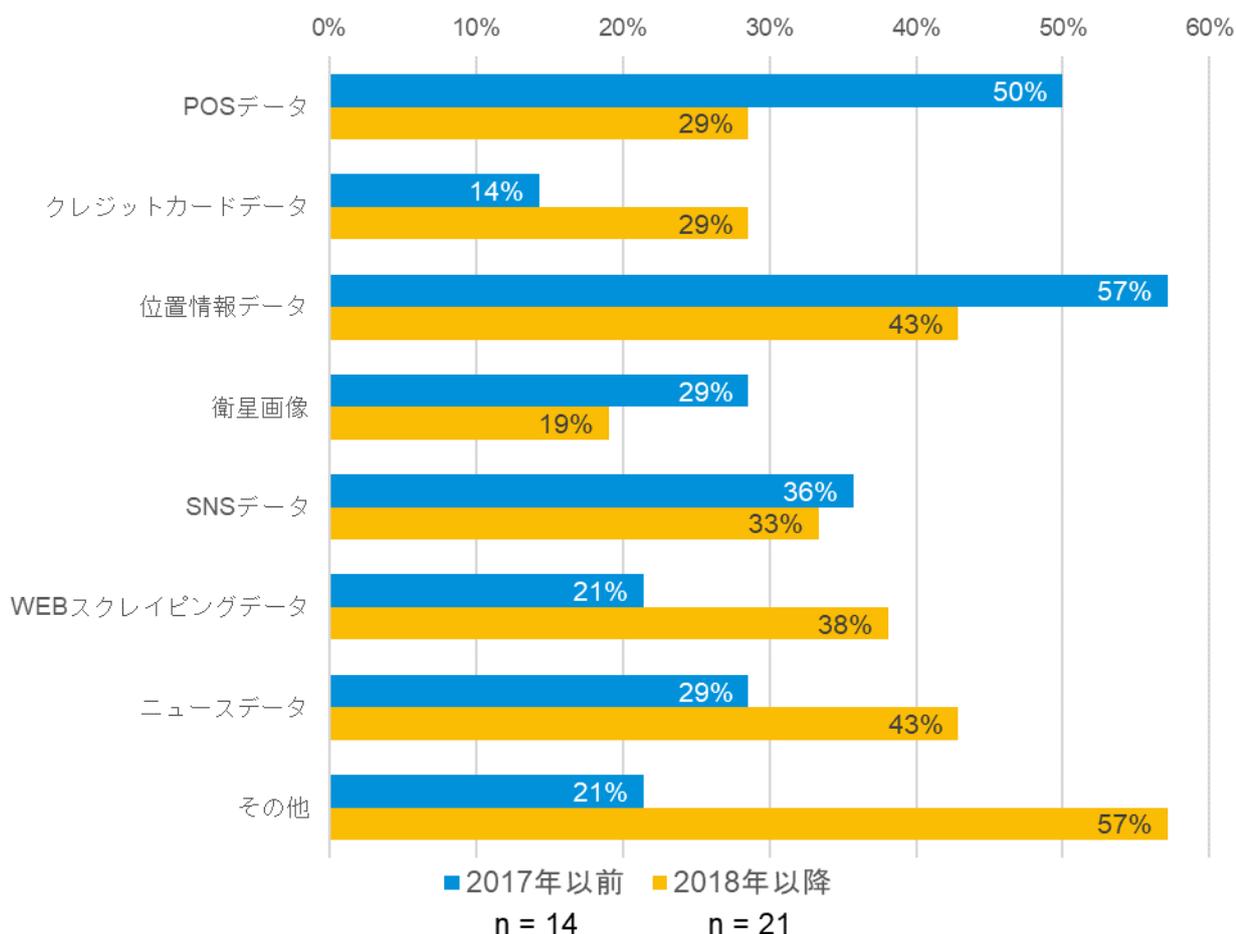
購入者・データプロバイダー・データ分析受託を対象に、2.(4)で集計したオルタナティブデータを利用開始した時期と、2.(9)の利用しているオルタナティブデータの種類の結果をもとに、クロス集計を行いました。

2017年以前からオルタナティブデータの利用を行っている層は、「位置情報データ(57%)」、「POSデータ(50%)」の利用率が高い傾向にありました。

一方で、2018年以降にオルタナティブデータを利用開始した層は、2017年以前の層と比較して「ニュースデータ(43%)」や「WEBスクレイピングデータ(38%)」の取扱いが多く、取扱い開始時期によって利用するデータの種類が異なることがわかりました。

また、「その他」の取扱いでは2017年以前の層が21%に対し、2018年以降の層は57%となっていることを踏まえると、2018年以降に参入した層は幅広い種類のオルタナティブデータを取扱っていることが示唆されます。

図表2-10 オルタナティブデータの取扱開始時期と扱っているデータの種類



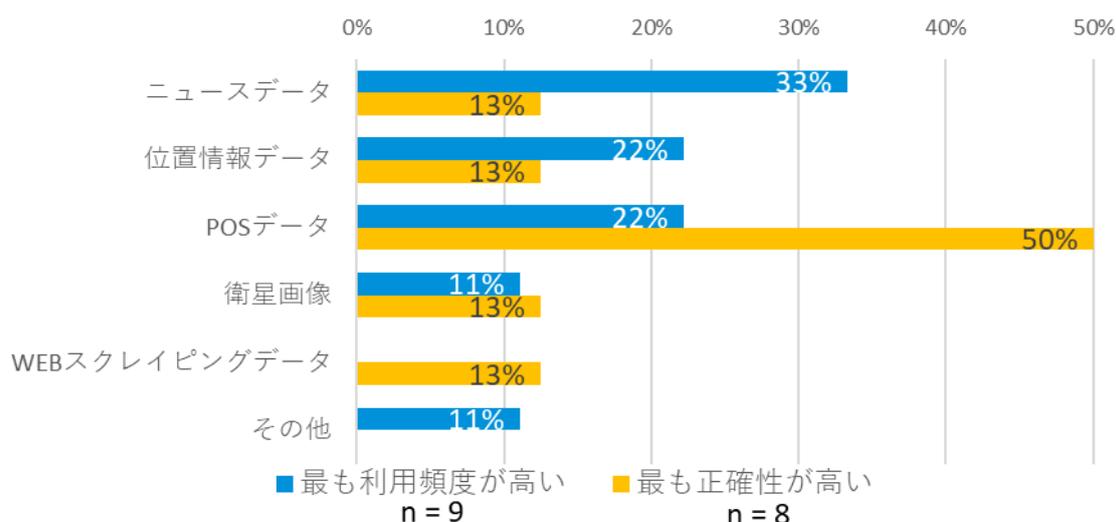
### (11) 最も利用頻度が高い／正確性が高いオルタナティブデータ

購入者に対して、オルタナティブデータのうち最も利用頻度が高いデータと、最も正確性が高いデータを質問しました。

最も利用頻度が高いデータは、「ニュースデータ（33%）」、「位置情報データ（22%）」、「POS データ（22%）」の順番となりました。

最も正確性が高いデータは、「POS データ（50%）」が最多となり、同率2位に「ニュースデータ」、「位置情報データ」、「衛星画像」、「WEB スクレイピングデータ<sup>1</sup>」が13%となりました。

図表2-11 最も利用頻度が高い／正確性が高いオルタナティブデータの種類の？

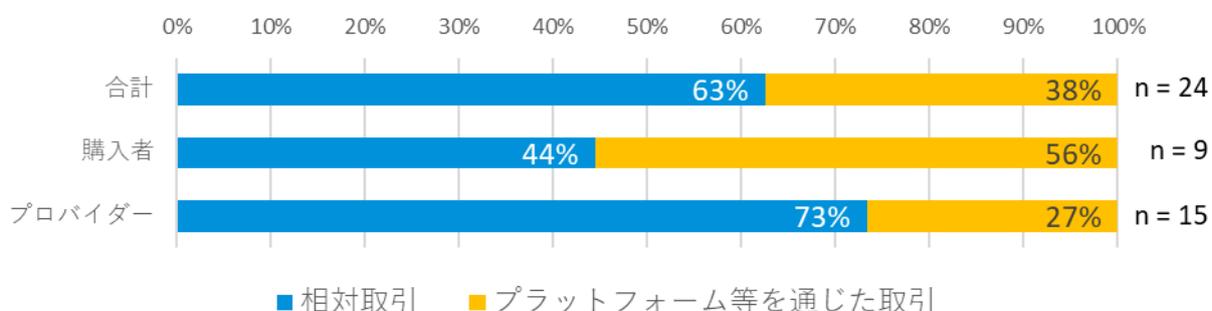


### (12) オルタナティブデータを売買しているチャネル

購入者とデータプロバイダーを対象に、オルタナティブデータを売買しているチャネルについて質問しました。

全体では、63%が相対取引を行っていることがわかりました。データプロバイダーに関しても73%が相対取引によるデータ売買を行っているのに対して、購入者はプラットフォーム等を経由した取引が過半の56%を占める結果となりました。

図表2-12 オルタナティブデータをどのチャネルで売買していますか？

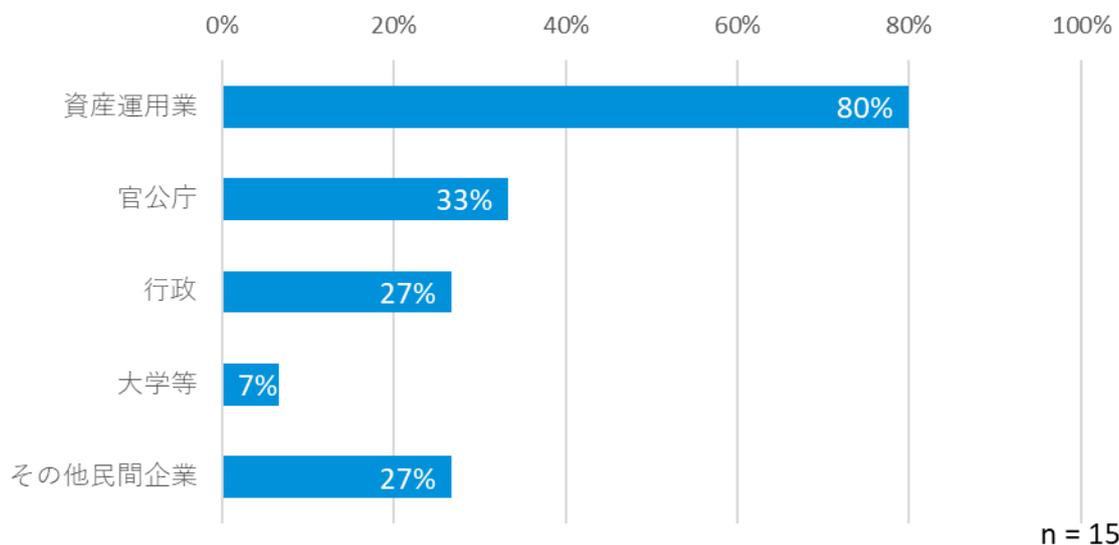


<sup>1</sup> Web上で公開されているページから、データ分析上有用な情報を自動収集されたデータ

### (13) オルタナティブデータの販売先

データプロバイダーを対象に、オルタナティブデータの販売先について質問しました。最も多かったのは「資産運用業向け（80%）」で、「官公庁（33%）」、「行政（27%）」、「その他民間企業（27%）」の順番になりました。

図表 2-13 オルタナティブデータの販売先は（複数選択可）？



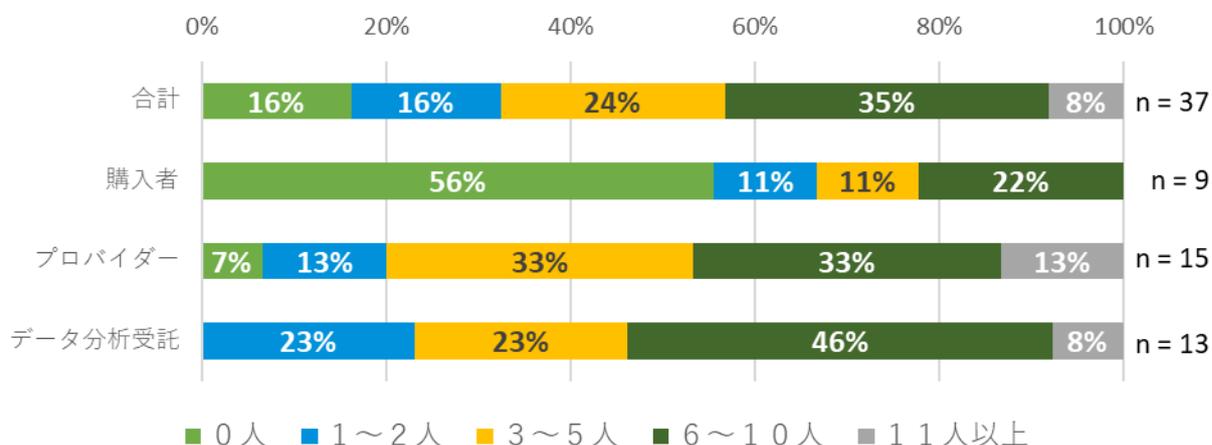
### 3. オルタナティブデータに関する人員／組織

#### (1) オルタナティブデータ専属の人員数

オルタナティブデータを扱う専属人員の数に対しては、全体では「6~10人（35%）」が最多となり、続いて「3~5人（24%）」の順番となりました。

主体別の内訳では、購入者では「0人（56%）」、「1~2人（11%）」と2人以下が過半を占めているのに対し、データプロバイダーや購入者では6人以上（データプロバイダー46%、データ分析受託 54%）の回答も目立ち、二極化の傾向がみられました。このことは、オルタナティブデータを利用する立場にある購入者を中心に体制整備が進んでいないことを示唆しています。

図表 3-1 オルタナティブデータ専属の人員は何人？

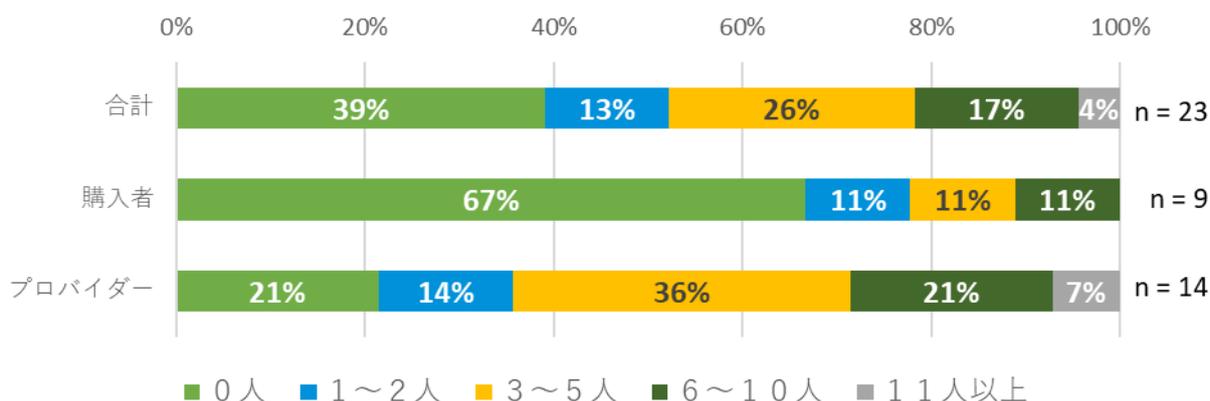


## (2) オルタナティブデータ専業のデータエンジニア／データサイエンティスト数

オルタナティブデータ専業のデータエンジニア／データサイエンティスト数を購入者とデータプロバイダーに確認しました。

購入者では、「0人(67%)」が最多となったのに対して、データプロバイダーでは「3~5人(36%)」、「6~10人(21%)」、「11人以上(7%)」との回答もあり、人員数に差がみられました。

図表3-2 オルタナティブデータ専業のデータエンジニア／データサイエンティストは何人？

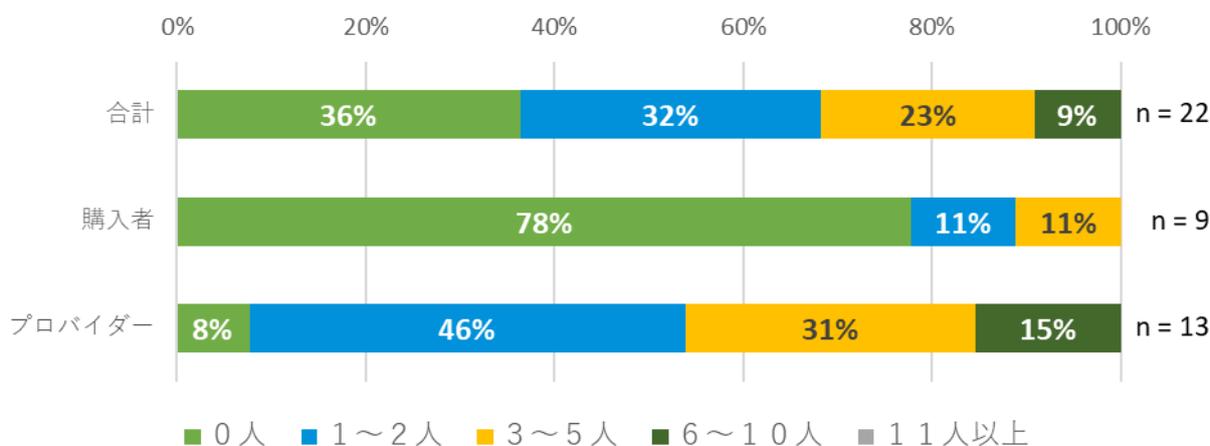


## (3) オルタナティブデータ専業のビジネスデベロップメントの人数

オルタナティブデータ専業のビジネスデベロップメント(事業開発)の人員数を購入者とデータプロバイダーに確認しました。

購入者では、「0人(78%)」が最多となったのに対して、データプロバイダーでは「1~2人(46%)」、「3~5人(31%)」、「6~10人(15%)」の順番となり、これまでの質問と同様に両者間の人員差がみられます。

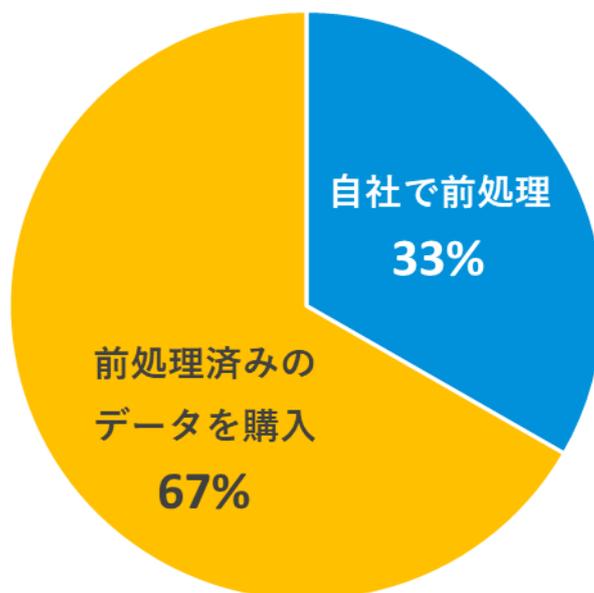
図表3-3 オルタナティブデータ専業のビジネスデベロップメントは何人？



#### (4) オルタナティブデータの前処理

購入者におけるオルタナティブデータの前処理については、前処理済みのデータを購入（66.7%）が多数を占めました。これは購入者でオルタナティブデータ専門の人材が少なかったこととも整合的です。

図表3-4 オルタナティブデータの前処理はどうしていますか？

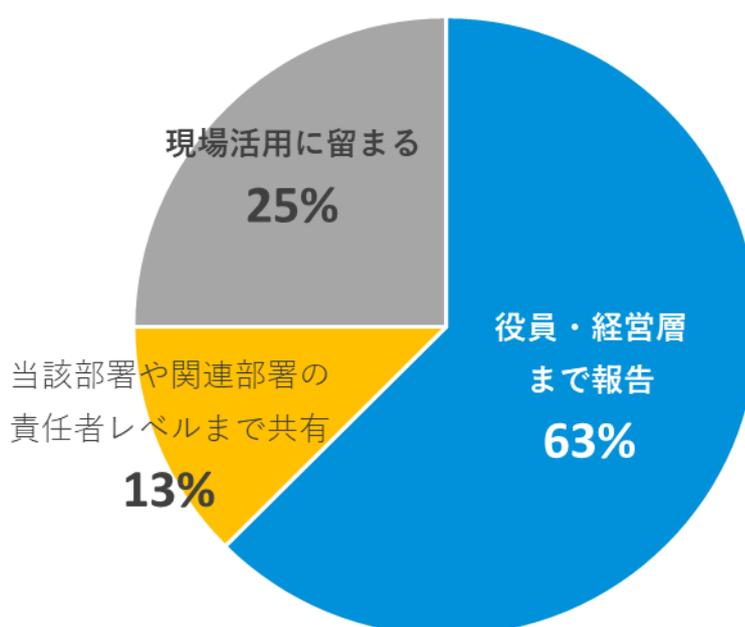


n = 9

#### (5) オルタナティブデータの社内共有方法

購入者に対して、オルタナティブデータを用いた分析結果の社内共有について確認したところ、「役員・経営層まで報告（63%）」が最多となりました。

図表3-5 オルタナティブデータやその分析結果の社内共有はどこまでしていますか？



n = 8

## 4. 法規制と課題

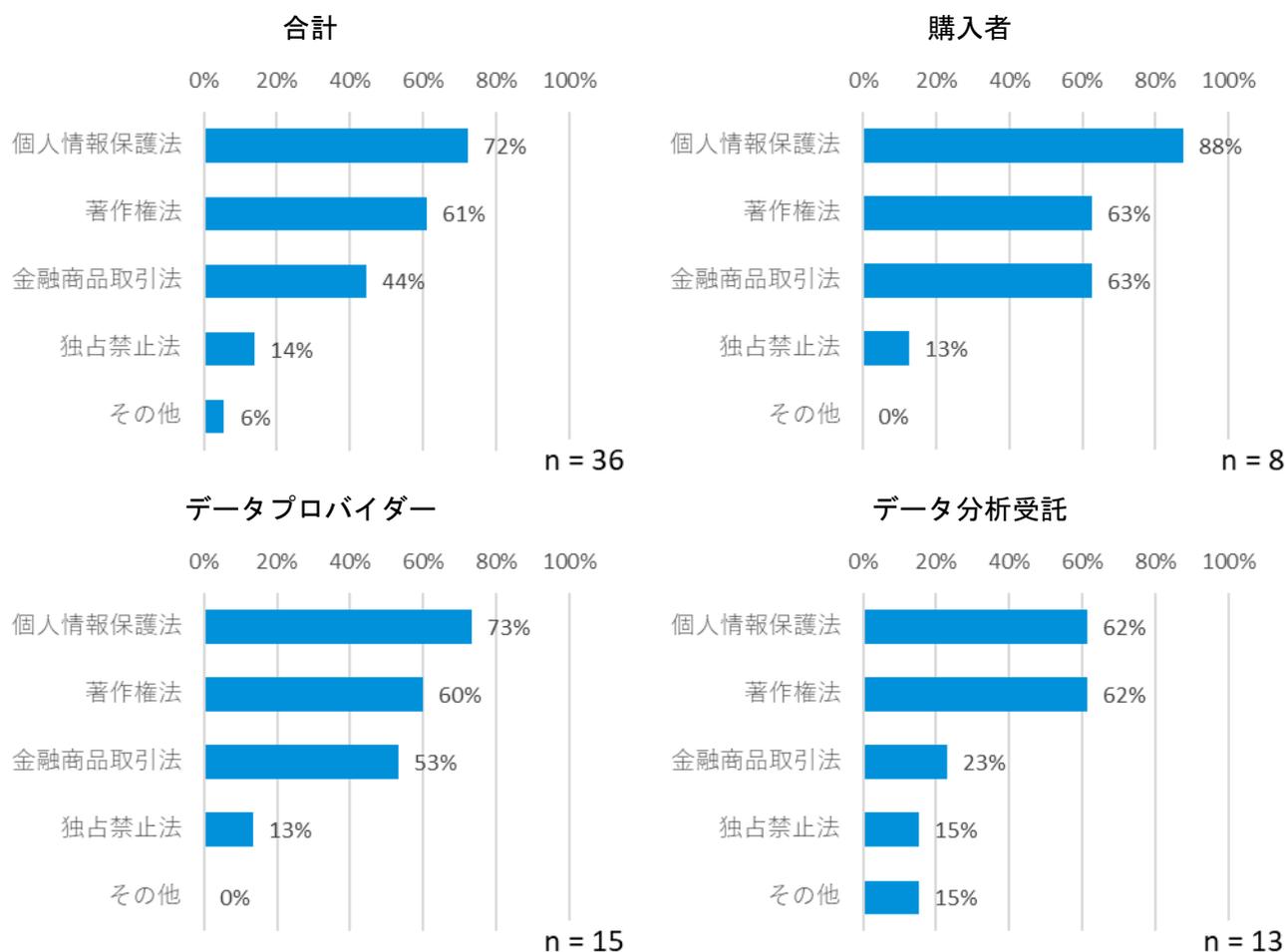
### (1) オルタナティブデータを利用するにあたって留意している法規制

オルタナティブデータを利用するにあたり留意している法律を、購入者、データプロバイダー、データ分析受託に質問しました。

全体では、「個人情報保護法(72%)」、「著作権法(61%)」、「金融商品取引法(44%)」、「独占禁止法(14%)」の順番となりました。

各主体で上記の順番は大きく変わりませんが、購入者が各法律に留意する割合が相対的に高くなっています。このことは、データ利用に際して注意を大きく払っている傾向にあることを示唆しています。また、購入者やデータプロバイダーに比較して、データ分析受託では「金融商品取引法(23%)」の回答が少ない結果となっています。

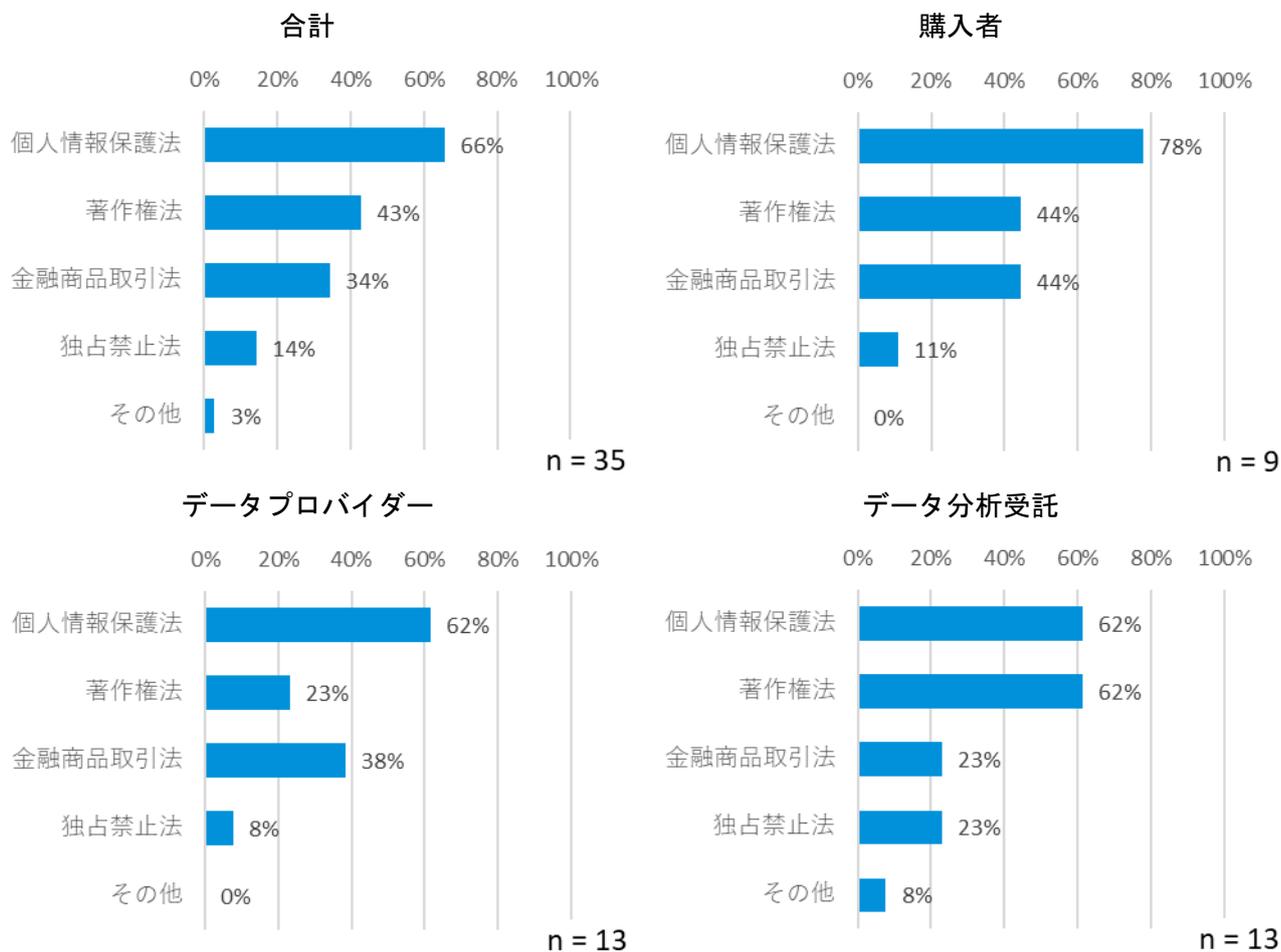
図表4-1 オルタナティブデータ利用するにあたって留意している法律は？（複数選択可）



## (2) オルタナティブデータを利用するにあたって整備が必要な法律

オルタナティブデータの利用にあたって、整備が必要な法律を質問したところに関しては、全体では「個人情報保護法（72%）」、「著作権法（61%）」、「金融商品取引法（44%）」、「独占禁止法（14%）」の順番となりました。

図表4-2 オルタナティブデータ利用するにあたって整備が必要な法律は？（複数選択可）



### (3) 取扱うオルタナティブデータの種類と法規制

購入者・データプロバイダー・データ分析受託を対象に、2.(9)で集計した取扱っているオルタナティブデータの種類(複数選択可)と、オルタナティブデータ利用時に留意している法律(複数選択可)でクロス集計を行いました。

個人情報保護法は、データの種類にかかわらず概ね留意が必要であるとの認識が共有されているようです。対して、金融商品取引法では、「ニュースデータ(67%)」や「POSデータ(50%)」の利用者で割合が高い一方、「衛星画像(13%)」の利用者においては割合が低くなっており、ばらつきが大きく見られました。

図表4-3-1 扱っているデータの種類の留意している法律

	個人情報保護法	金融商品取引法	著作権法	独占禁止法	その他	n
POS	82%	55%	73%	18%	9%	11
クレジットカード	83%	33%	50%	0%	0%	6
位置情報	88%	44%	69%	6%	0%	16
衛星画像	88%	13%	75%	0%	13%	8
SNS	70%	50%	60%	0%	20%	10
WEBスクレイピング	70%	30%	70%	10%	20%	10
ニュース	67%	67%	75%	8%	0%	12

同様に、購入者・データプロバイダー・データ分析受託を対象に、2.(9)で集計した取扱っているオルタナティブデータの種類(複数選択可)と、オルタナティブデータ利用時に整備が必要な法律(複数選択可)でクロス集計を行いました。

個人情報保護法は、総じて法律の整備が必要だとする割合が高い結果となりました。一方、金融商品取引法、著作権法、独占禁止法では、大きなばらつきが見られました。これらの結果を踏まえると、扱うデータによって対象となる法律がまちまちであり、そのことが利活用を妨げている可能性があります。

図表4-3-2 扱っているデータの種類の整備が必要な法律

	個人情報保護法	金融商品取引法	著作権法	独占禁止法	その他	n
POS	73%	45%	45%	9%	0%	11
クレジットカード	86%	43%	71%	14%	0%	7
位置情報	80%	27%	47%	0%	0%	15
衛星画像	86%	0%	86%	0%	0%	7
SNS	70%	50%	50%	20%	10%	10
WEBスクレイピング	90%	40%	70%	30%	10%	10
ニュース	73%	64%	64%	27%	0%	11

#### (4) オルタナティブデータを活用するにあたっての課題

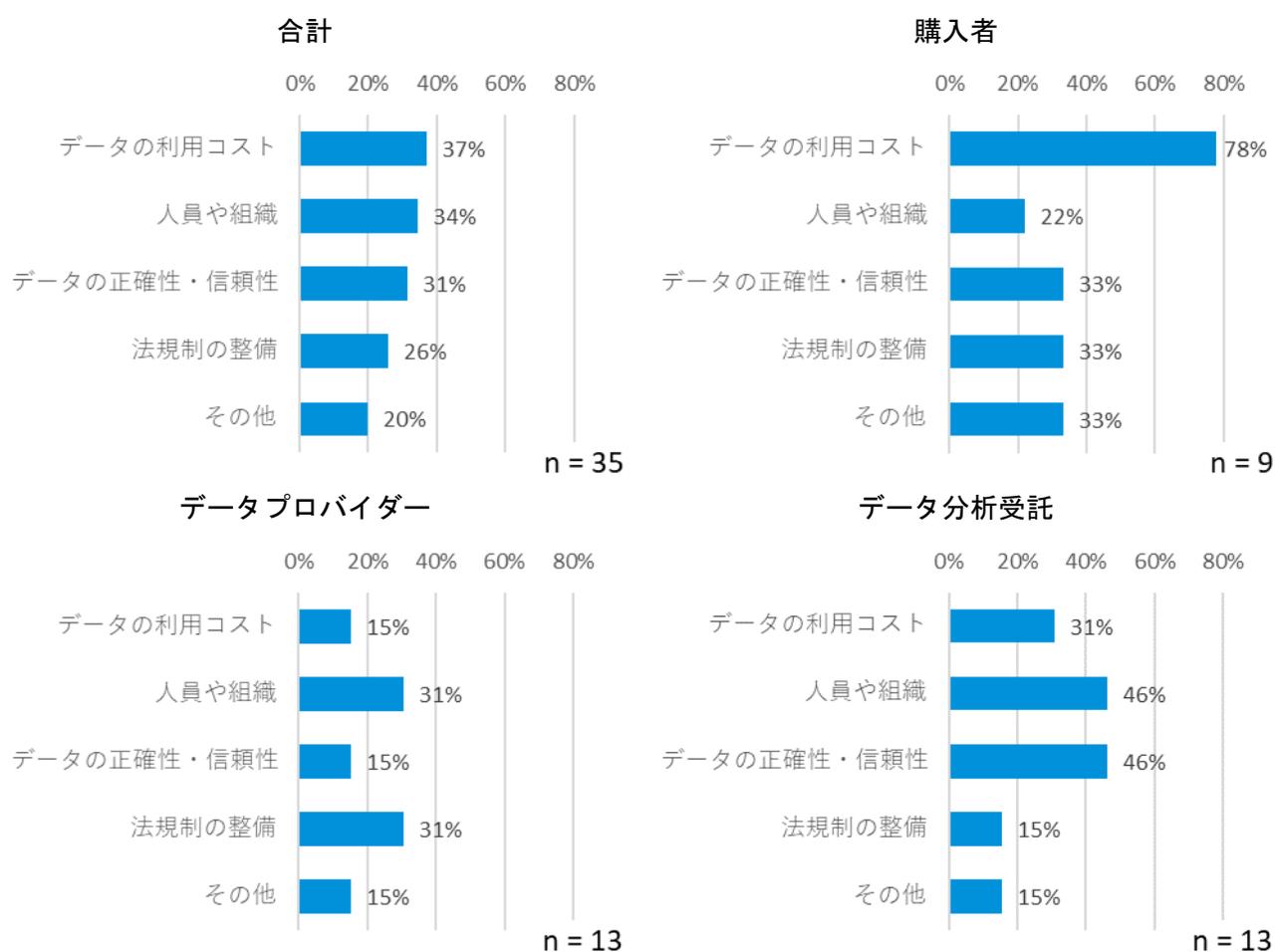
オルタナティブデータを活用するにあたっての課題について、全体では「データの利用コストが高い (37%)」、「データを取扱う人員がない (34%)」、「データの正確性、信頼性が担保されない (31%)」、「法規制が整備されていない (26%)」の並びとなりました。

購入者では、「データの利用コストが高い (78%)」がほかの項目を突き放して最多となりました。

データプロバイダーでは、「データを取扱う人員がない (31%)」、「法規制が整備されていない (31%)」が同率となりました。

データ分析受託では、「データを取扱う人員がない (46%)」、「データの正確性、信頼性が担保されない (46%)」が最多で並び、それぞれで違う課題を抱えていることが明らかになりました。

図表 4-4 オルタナティブデータを利用するにあたっての課題は？ (複数選択可)



また、自由記述形式の課題については下記のような意見が寄せられました。

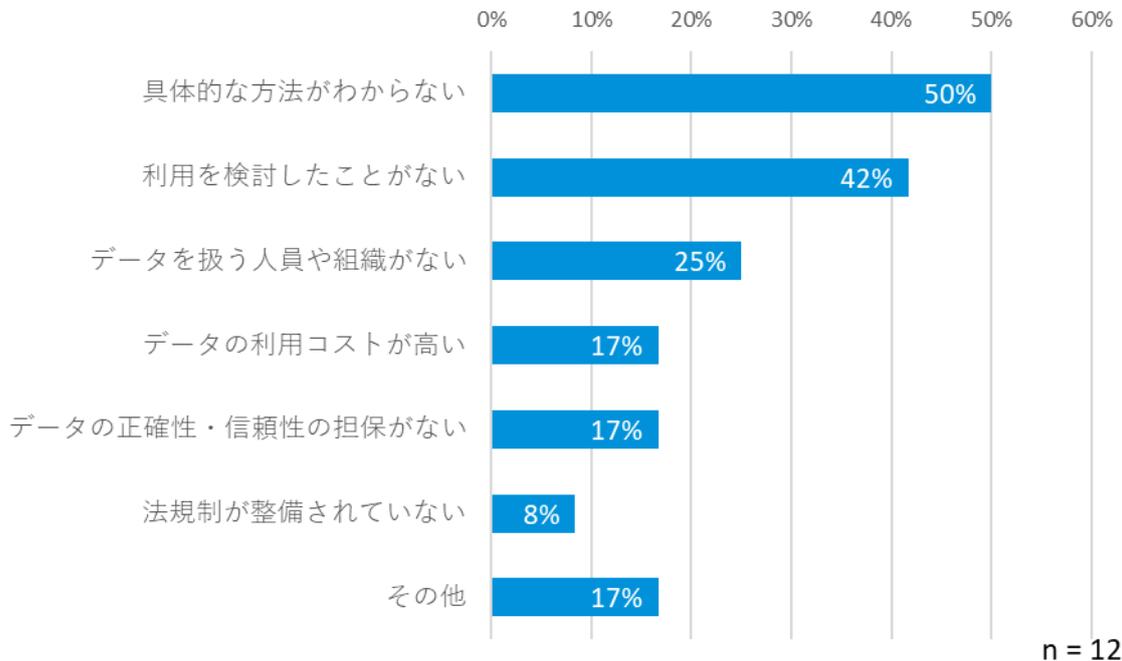
購入者	データを購入・トライアルしてみると、先方の担当者も気づかなかった不具合が見つかる場合が多い。
	複数のオルタナティブデータをかけあわせることで価値を生むことが多いものの、複数を買うことは予算的に困難。
	分析を行うためのツールやインフラ整備のコストが高い。
	サンプリングバイアスの程度、過去の長期データが不足している。
	各種データの利用方法、利用目的に関する理解が社内で確立されておらず、購入するデータの内容や費用に関する共通理解を得ることが困難。
データプロバイダー	レセプトデータや健康診断データは国家での法整備がされておらず、個人情報保護法や業界団体からの圧力があり、世の中に出せないデータが数多くある。
	金融商品取引法など、現行法に抵触するかもしれないかが明確ではないため、販売やサービス拡大をする際、必要以上に慎重にならざるを得ず、オルタナティブデータの普及や利用推進に影響がある。
	弊社もそうですけれどもアプリから取得することで生じるアプリのユーザー特性や偏りを正しく説明し活かせるようにする。
	マーケットを拡大させるためには、データを扱える人員が増えることが必須。
	センシングに使用するセンサーやスマートフォン等も変化するため、蓄積されたビッグデータから均質にデータを取得（サンプリング）する方法を常に検証し続ける必要がある。（統計的な断層の問題）
	海外では理解があるが、日本ではオルタナティブデータの活用が難しい。
	日本国内企業にオルタナティブデータを扱える人材や組織があまりないので、オルタナティブデータを販売する側のコストが高くなりがちである。
データ分析受託	海外企業への販路がない。日本企業はWEBオルタナデータに関しての購入に関してはまだ慎重。仮に購入出来ても安価である可能性が大。
	極めて大容量のデータに対するデータマネジメント、クラウド環境でのコストマネジメント。
	オルタナティブデータをビジネスにどのように役立てるか、というノウハウの共有がより必要。
	共通ポイントカードの会員が日本全体を表しているのかなどデータの信頼性は、判断しづらい。また、分析環境を準備するのに時間が掛かり、スピーディーな対応ができない。
	分析のご利用を検討しているお客様が具体的なユースケースを思いつきにくいいため、ニーズを聞いた上で類似のユースケースを示す必要があります。
	発注側にデータを取扱う専門家が少ない。
	衛星データのコストが高く、活用が進まない。今後衛星機数が増えるに従って単価は下がっていくと予測。活用を拡大するためにも衛星データの精度や利用方法についての専門家でなくてもわかりやすい基準づくりが必要。
	新しい領域のため、リクエストに対応するのに専門的なリソース不足が発生する。

## 5. 未利用者

### (1) オルタナティブデータを利用していない理由

未利用者がオルタナティブデータを利用していない理由は、「具体的な方法がわからない (50%)」、「利用を検討したことがない (42%)」、「データを扱う人員や組織がない (25%)」の順となりました。

図表5-1 オルタナティブデータを利用していない理由は？（複数選択可）



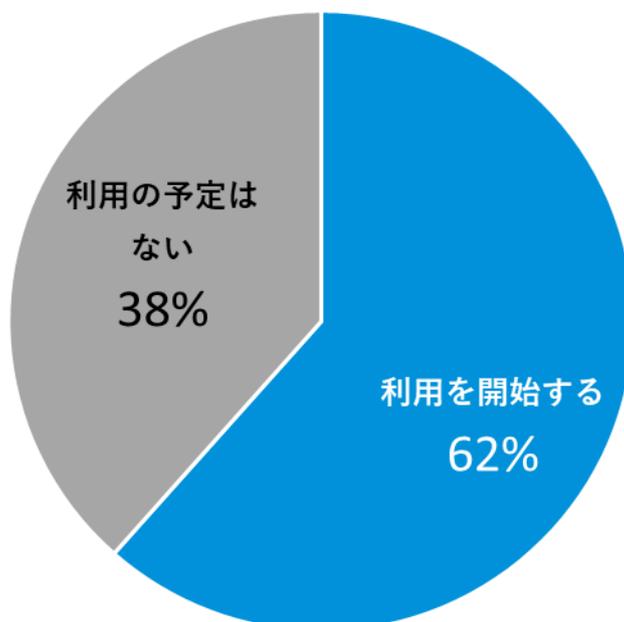
また、自由記述ではオルタナティブデータを利用していない理由について、以下のような回答が寄せられました。

ユーザー、プロバイダ共に明確な活用事例がないと利用に踏み切れない国民性がある。
マーケットの予測モデルにオルタナティブデータの利用を検討したが、期待する精度向上につながらなかった。
人流データ等を用いて商業施設などの価値評価への応用を考えているが、簡易、安価で利用できる方法がよくわからない。
オルタナティブデータを利用する立場としては、弊社が所有しているデータとオルタナティブデータを組み合わせたビジネスモデルを模索中。オルタナティブデータを販売する立場としては、弊社からオルタナティブデータとして販売できるデータまで至っていない。
研究・教育用の予算でまかなえるようなデータベースの情報があるとよい。
お客さま自身がデータ分析に価値を見出す動機付けがまだできていない。データ分析＝回答を持って来てくれる構図から、協業でどうマネタイズするかに変えていくところに課題がある。
オルタナティブデータがどのようなものか理解してない。
オルタナティブデータの認知度の向上、法規制の整備が必須。

## (2) 今後のオルタナティブデータ利用方針

未利用者に対して、今後（今年度含む3年程度）のオルタナティブデータ利用方針を質問したところ、「利用開始する」が62%を占めました。

図表5-2 今後（今年度含む3年程度）のオルタナティブデータ利用方針は？



n = 13

## 6. まとめ

本稿では多様な質問を投げかけ、これまで明らかになっていない国内のオルタナティブデータの活用状況を定量的に評価しようと試みました。結果的に何点か示唆が得られましたので、下記に紹介します。

第一に、「オルタナティブデータの利用が足元で拡大している」という言説が、アンケートによって裏付けられました。オルタナティブデータをすでに利用している企業のうち、2021年度の取扱額が前年比で「大幅に増えた」、「増えた」と回答が68%を占め、残りは「横ばい」となっています。また、ここ数年でオルタナティブデータを取扱い始めたプレイヤーも多く存在し、今後3年間で取扱額を拡大するとした回答が大半を占めています。これらを総合すると、国内のオルタナティブデータ市場は今後も拡大していくことが見込まれます。

第二に、オルタナティブデータには数多くの種類が存在し、その利用価値は誰が利用するか、どのように利用するかによって、評価が変わることがわかりました。位置情報データは、購入者によって最も取扱いされている割合が高いデータでした。一方で、購入者へ最も利用頻度が高い／正確性が高いオルタナティブデータを質問したところ、それぞれニュースデータとPOSデータが最多となっています。これらの結果は、個々のオルタナティブデータが、長所や短所など異なる特性を持っているということを示唆しています。オルタナティブデータの歴史は浅く、そのバイアスや分析手法も発展途上にあります。今後は利用方法や目的に応じて、どのようなオルタナティブデータを用いるべきかといった知見の蓄積が期待されます。

第三に、オルタナティブデータの利用を拡大するためには、解決すべき種々の課題があることが明らかになりました。今回のアンケートでは、「データの利用コストが高い」、「データを扱う人員や組織がない」、「データの正確性・信頼性が担保されない」、「法規制が整備されていない」などの課題が主に挙げられました。

オルタナティブデータの利用コストについて、データそのものの価格に加えて、それに付随する分析環境などの間接的な費用についても言及がありました。また、これらのコストが高いという事実に加えて、「購入するデータの内容や費用に関する共通理解を得ることが困難」といったコメントも寄せられています。オルタナティブデータの活用実績が蓄積し、その便益が広く明らかになれば、費用の割高感が解消される可能性があります。また、特に分析費用については、今後のオルタナティブデータ普及とともに逡減していくと考えられます。

オルタナティブデータを扱う人員や組織がないという課題は、購入者・データプロバイダー・データ分析受託で共通の悩みようです。特に、購入者ではデータプロバイダーやデータ分析受託と比較して、オルタナティブデータに特化した人材が少なくなっています。分析を行うスキルが不足していることに加え、オルタナティブデータからどのような知見が得られるかがイメージできないといった潜在的な問題もあります。データ分析スキル向上だけでなく、幅広くユースケースを共有する場が重要だと考えられ、JADAAが毎週提供している勉強会の場は有効であると考えられます。

データの正確性・信頼性が担保されないといった課題も多く指摘されました。オルタナティブデータはその歴史の浅さから長期データが存在せず、どのようなバイアスを内包しているのかは研究途上にあります。オルタナティブデータは主に民間企業の経済活動に付随する副産物であるため、社会状況を把握するために設計された公的統計と比べて、バイアスが強く含まれる傾向にあります。したがって、オルタナティブデータにはバイアスがあることを認識したうえで、バイアスを最小化するためにどのような

処理を行うべきかというアプローチが必要です。利用・研究を深めてバイアスを認識し、それをユーザー共通の知見とすることによって、中長期的にこうした課題は解決の方向に向かっていくでしょう。

法規制が整備されていないという問題によって、オルタナティブデータに対して慎重になっている企業も散見されました。データ活用と個人情報保護はトレードオフになりがちであり、多くの回答で個人情報保護法に留意し、今後の整備が必要だという回答が見られました。また、「現行法に抵触するかしな  
いか  
が明確ではないため、販売やサービス拡大をする際、必要以上に慎重にならざるを得ず、オルタナティブデータの普及や利用推進に影響がある」といった意見もありました。JADAA では、データ提供者に対し法令違反及びその他リスクの検出を効果的に検証できる確認事項（DDQ）を公開しており、法整備の提言やガイドマップの策定などを進めています。

先に述べたように、今後もオルタナティブデータの活用は進展していくと見込まれています。オルタナティブデータを活用するメリットは周知されつつありますが、種々の課題の解消に向けた取り組みも同時に進めていかなければなりません。オルタナティブデータの活用状況や課題を把握するために、本FACTBOOK は今後も継続して発行していく予定です。

委託者：一般社団法人オルタナティブデータ推進協議会

受託者：SOMPO インスティテュート・プラス株式会社

JADAA FACTBOOK プロジェクト：赤井 厚雄、東海林 正賢、辻中 仁士、小島 治樹