



JADAA

オルタナティブデータ推進協議会

# オルタナティブデータ FACTBOOK2024年度版 (概要)

～支出額の増加ペースは鈍化も、データ活用の多様化や人員拡充がみられる～

2024年12月

一般社団法人オルタナティブデータ推進協議会

Japan Alternative Data Accelerator Association

SOMPOインスティテュート・プラス株式会社

Sompo Institute Plus Inc.

小池理人

菅沼健司



# 本調査の全体像

## 問題意識

- 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、利用実態に関する調査研究も進展している。
- 日本では「どのようなオルタナティブデータが」、「どのような目的で」、「どの程度利用されているか」といった実情が明らかになっていない。国内におけるオルタナティブデータの現状や課題を把握することで、利活用がいつそう進展するのではないか。

## 本調査について

- 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、オルタナティブデータ推進協議会の会員に対してアンケートを実施した（2022年開始、今回は第3回）。
- 全体で56先からの回答を得て、FACTBOOKとして取りまとめを行った。

### ■ 本調査の回答者

業種	回答数	割合
資産運用業	5	9%
金融保険業（除く資産運用業）	7	13%
メーカー	2	4%
シンクタンク	3	5%
教育機関	7	13%
IT/システム開発	21	38%
その他	11	20%
総計	56	100%

### ■ 本調査の集計期間

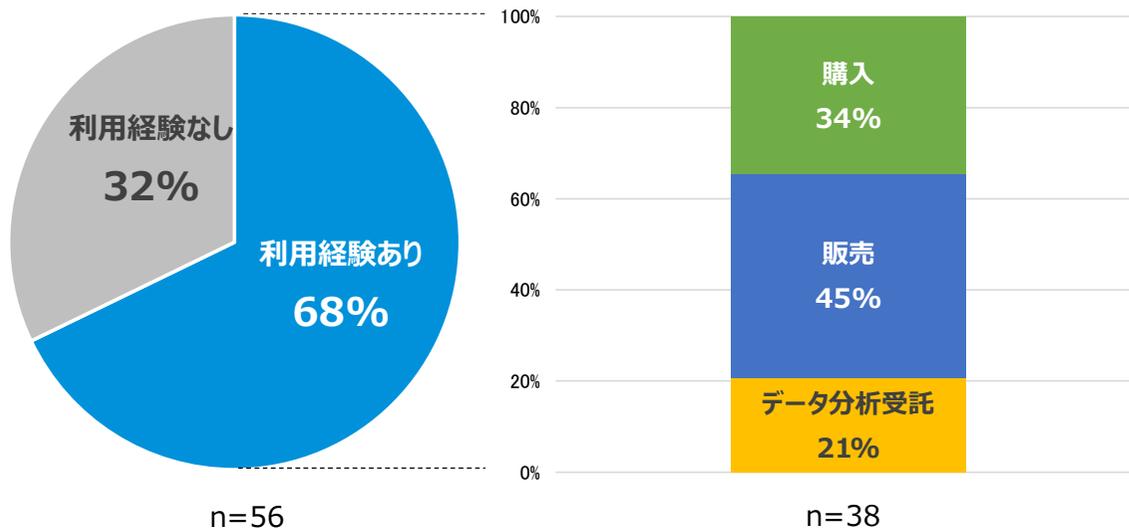
2024年8月15日～2024年10月15日

# オルタナティブデータの利用経験

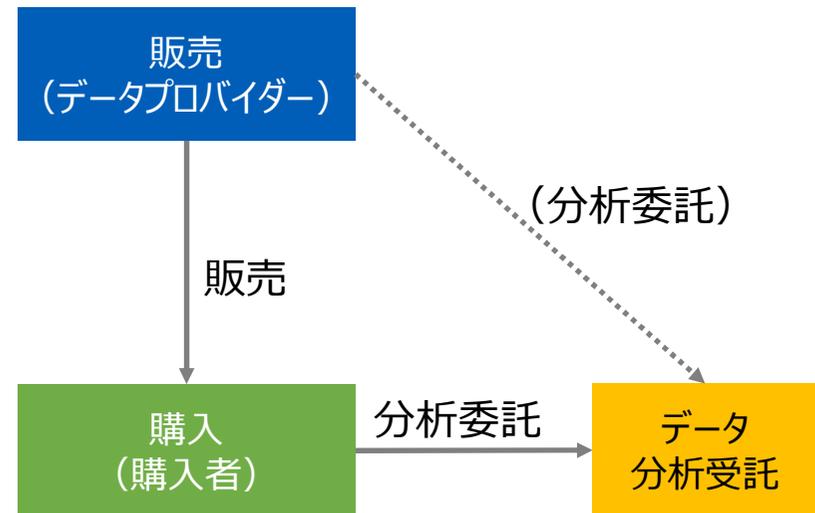
購入者	データプロバイダー
データ分析受託	未経験

- オルタナティブデータの利用経験を質問したところ、68%が「利用経験あり」と回答。
- 「経験あり」とした回答の内訳は、「購入経験あり」が34%、「販売経験あり」が45%、「データ分析受託経験あり」が21%となった。
- これらの各主体を「購入者」「データプロバイダー」「データ分析受託」の3つに分類し、以降はそれぞれに質問を行った。

## ■ オルタナティブデータの利用経験があるか？



## ■ 各主体のイメージ

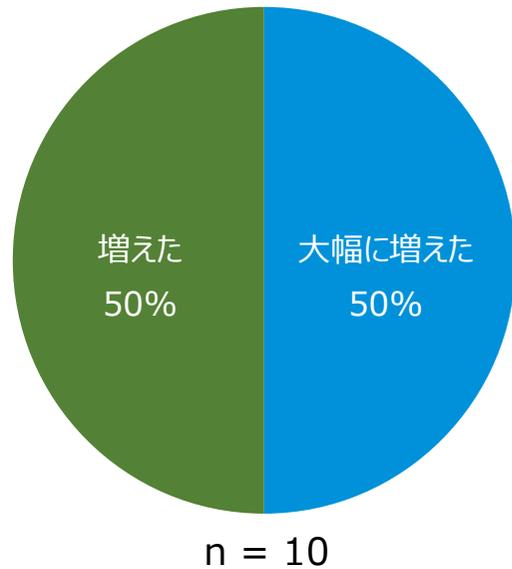


# 支出金額の増減

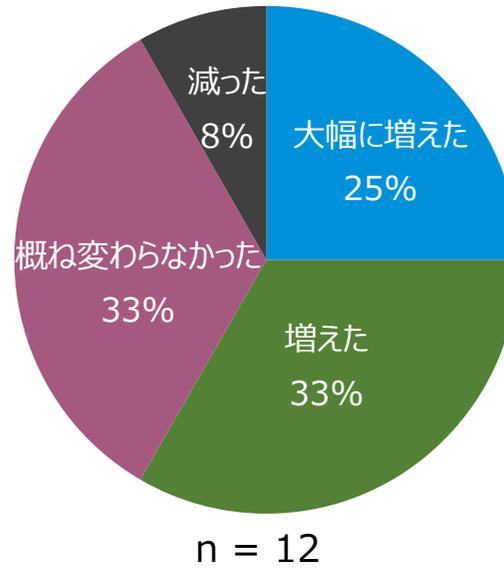
- 購入者に対し、オルタナティブデータへの支出金額の前年度からの変化を質問した。
- 「概ね変わらなかった」が85%と大部分を占め、「増えた」は15%に止まった。
- オルタナティブデータへの支出額については、増加ペースが次第に鈍化しつつある傾向がみられる。

## ■ オルタナティブデータの取扱額は前年度と比べてどう変化したか

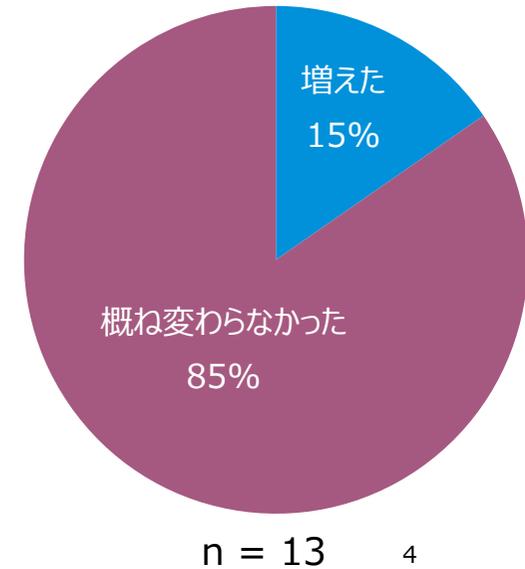
2022年調査



2023年調査



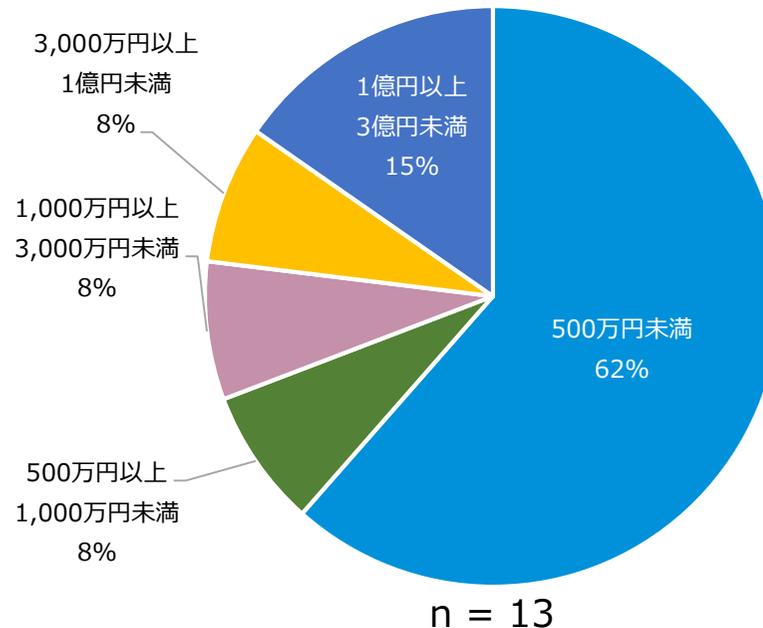
2024年調査



# 年間の支出金額の分布

- 購入者に対し、オルタナティブデータの年間支出金額をレンジで質問した。
- 「500万円未満」が62%を占め、「1億円以上3億円未満」が15%で続き、それ以外の回答は少数に止まった。
- オルタナティブデータの単価などを踏まえると、大々的にデータを活用している先は少なく、多くの先で活用は小規模に止まっていることが窺われた。

## ■ オルタナティブデータに関する年間の支出金額はいくらか

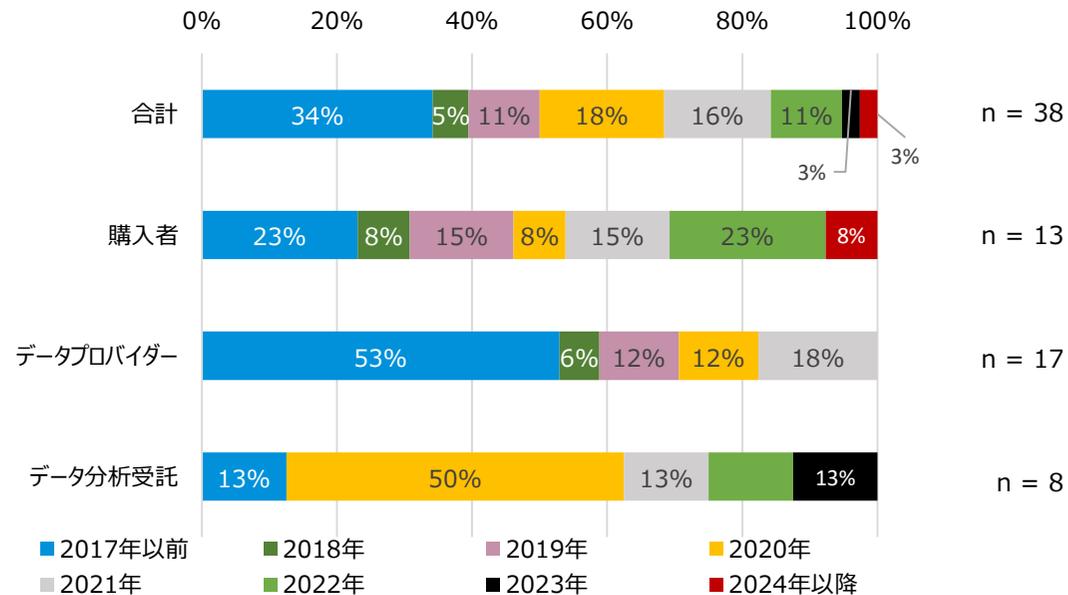


# 利用開始時期

購入者      データプロバイダー  
データ分析受託

- オルタナティブデータの利用開始時期については、全体では「2017年以前」が最多となった。他方で2020年以降に利用を開始した先も多く、コロナ禍を背景に需要が拡大したことが示唆される。
- 各主体の内訳も上記の傾向と一致しているが、データプロバイダーやデータ分析受託では利用開始時期が早い一方、購入者は2022年以降に利用開始した先も相応に存在している。

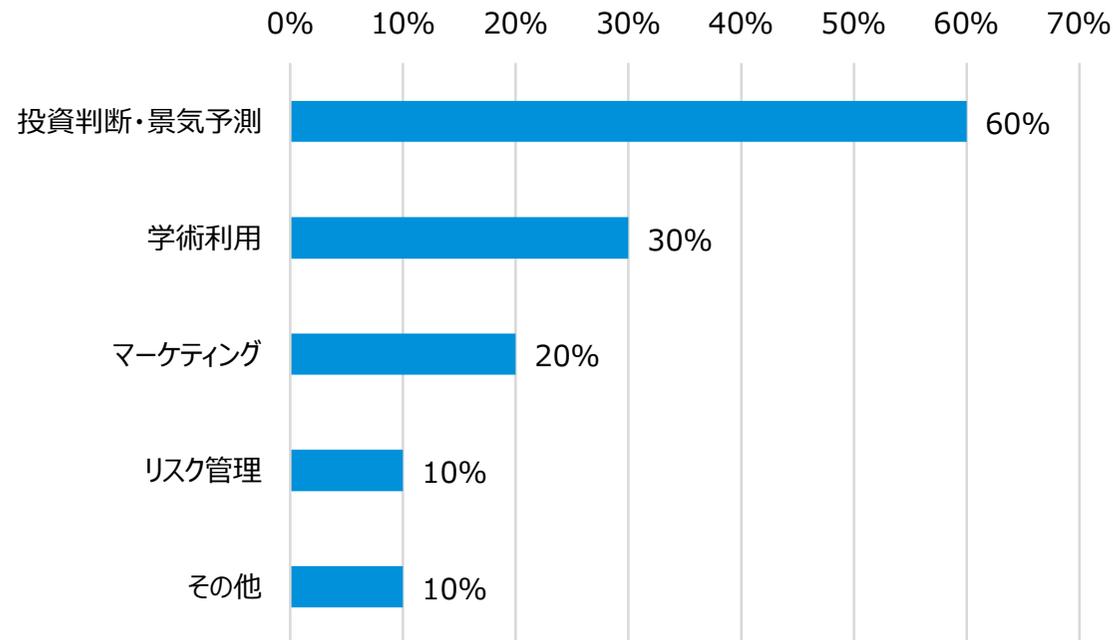
## ■ オルタナティブデータを利用開始した時期はいつか



# オルタナティブデータの活用目的

- 購入者に対し、オルタナティブデータの活用目的を質問した。
- 活用目的としては、「投資判断・景気予測」が最多となった。その他には、「学術利用」「マーケティング」「リスク管理」が続いた。

## ■ オルタナティブデータの活用目的は何か（複数回答可、最大3つまで）

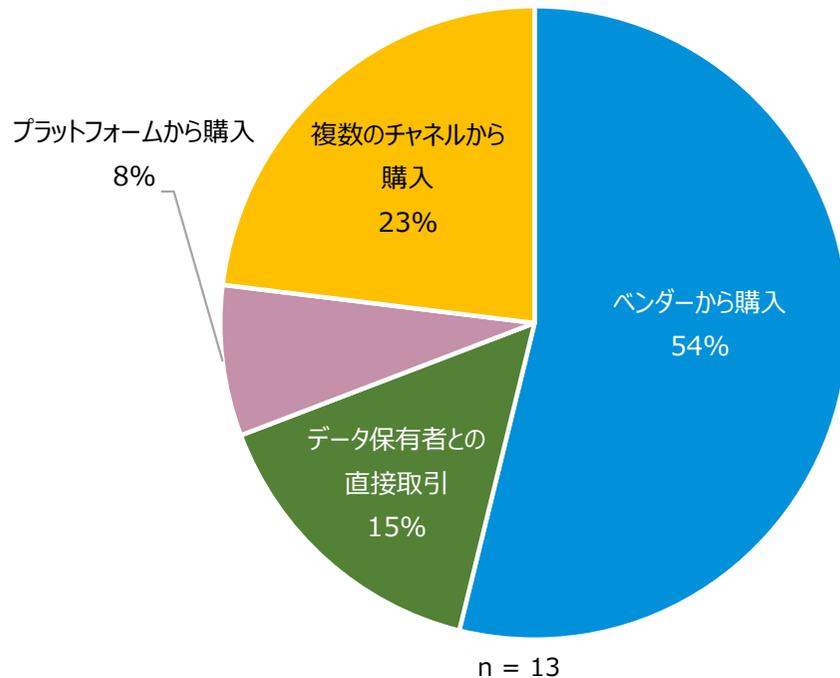


n = 10

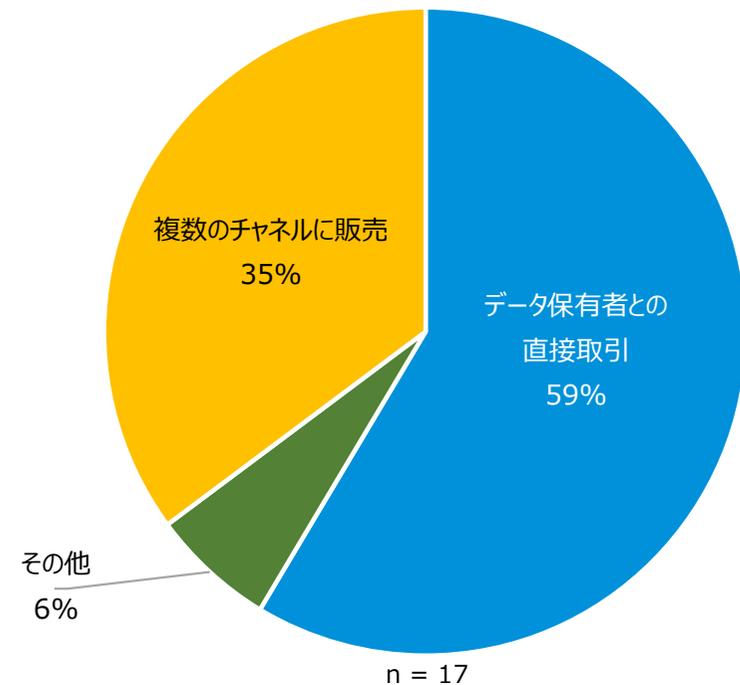
# オルタナティブデータの取引チャンネル

- オルタナティブデータの購入（販売）しているチャンネルについては、購入者からは「ベンダーから購入」、プロバイダーからは「データ保有者との直接取引」との回答が最も多く聞かれた。
- 一方、「プラットフォームからの購入」は極めて少ない回答に止まった。

## ■ オルタナティブデータを購入しているチャンネルは何か



## ■ オルタナティブデータを販売しているチャンネルは何か



# オルタナティブデータの利点

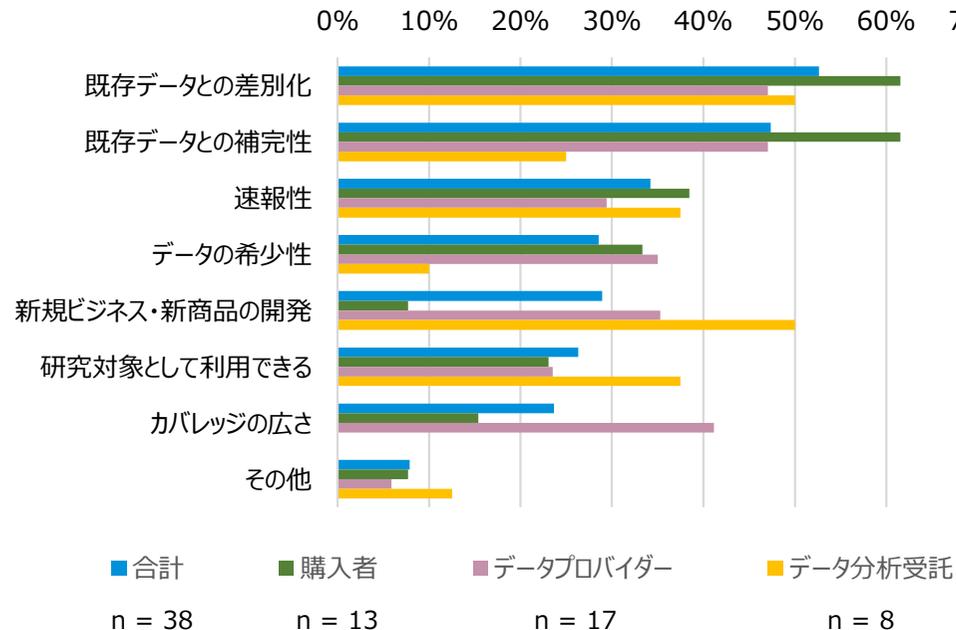
購入者

データプロバイダー

データ分析受託

- オルタナティブデータを利用する利点については、全体では「既存データとの差別化」、「既存データとの補完性」、「速報性」といった回答が多くみられた。
- 内訳をみると、購入者では「既存データとの差別化」や「既存データとの補完性」を重視する一方、データプロバイダーでは「データの希少性」が、データ分析受託では「新規ビジネス・新商品の開発」を重視する結果が示された。

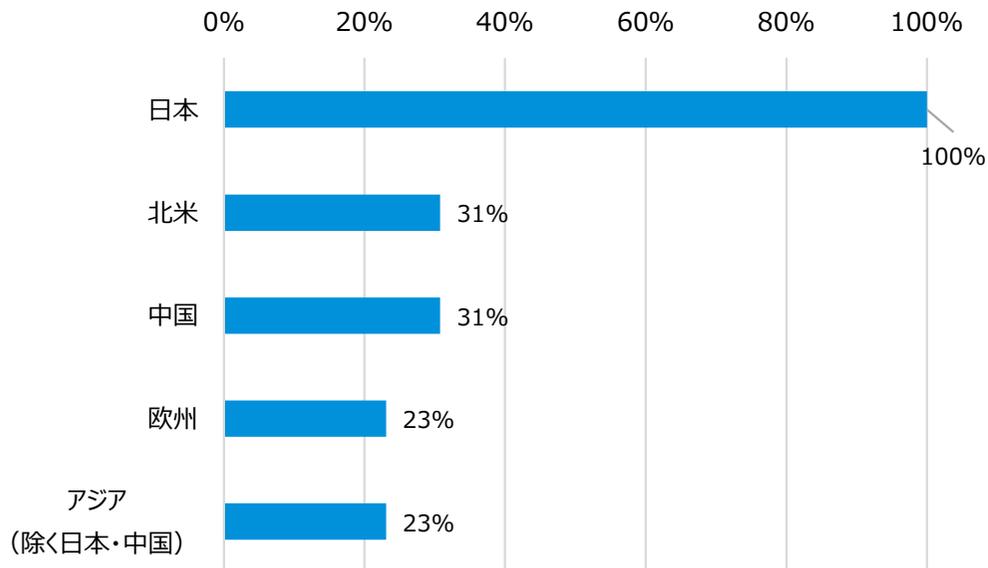
## ■ オルタナティブデータ利用の利点は何か（複数回答可、最大3つまで）



# 対象地域／対象セクター

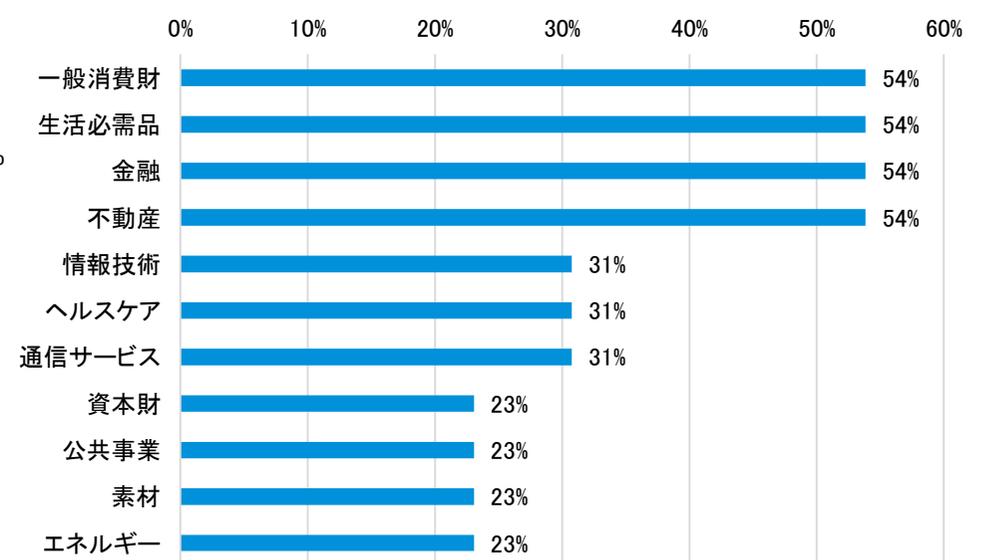
- 購入者に対して、購入しているオルタナティブデータの対象地域を確認したところ、「日本」が最多となった。海外では、「北米」「欧州」「中国」「日本と中国を除くアジア」とも、一定の利用がみられた。
- 同じく対象セクターに対する質問では、「一般消費財」「生活必需品」「金融」が多く挙げられた。「不動産」「情報技術」「ヘルスケア」「通信サービス」などの利用も多く聞かれた。

■ 購入データの対象地域（複数回答可）



n = 13

■ 購入データの対象セクター（複数回答可）



n = 13

# 取扱っているオルタナティブデータの種類

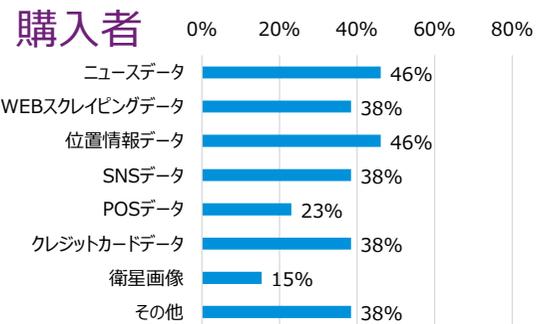
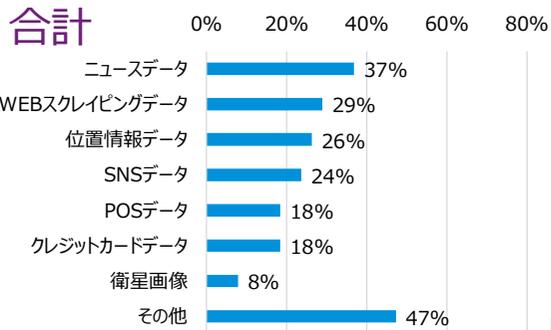
購入者  
データ分析受託

データプロバイダー

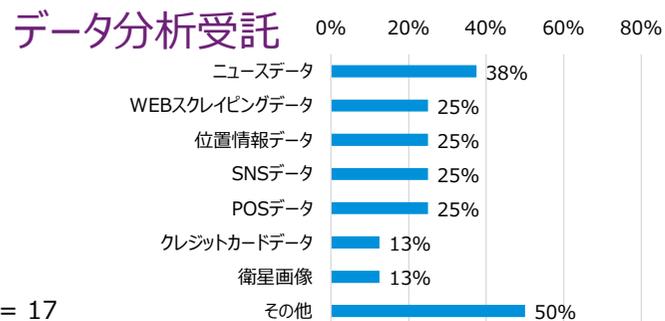
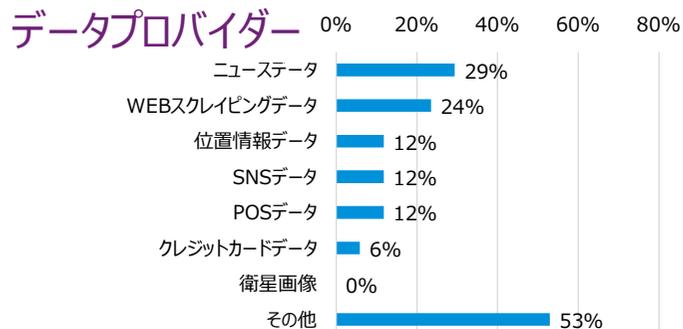
- 取扱っているオルタナティブデータの種類については、全体では「ニュース」「WEBスクレイピング」「位置情報」などバランス良くデータが取り扱われていることが示された。
- データプロバイダーやデータ分析受託における「その他」をみると、「口コミデータ」「医療データ」「テレビ視聴率データ」など多様なデータの利用が行われている。

## 取扱っているオルタナティブデータの種類（複数回答可）

## ■ その他



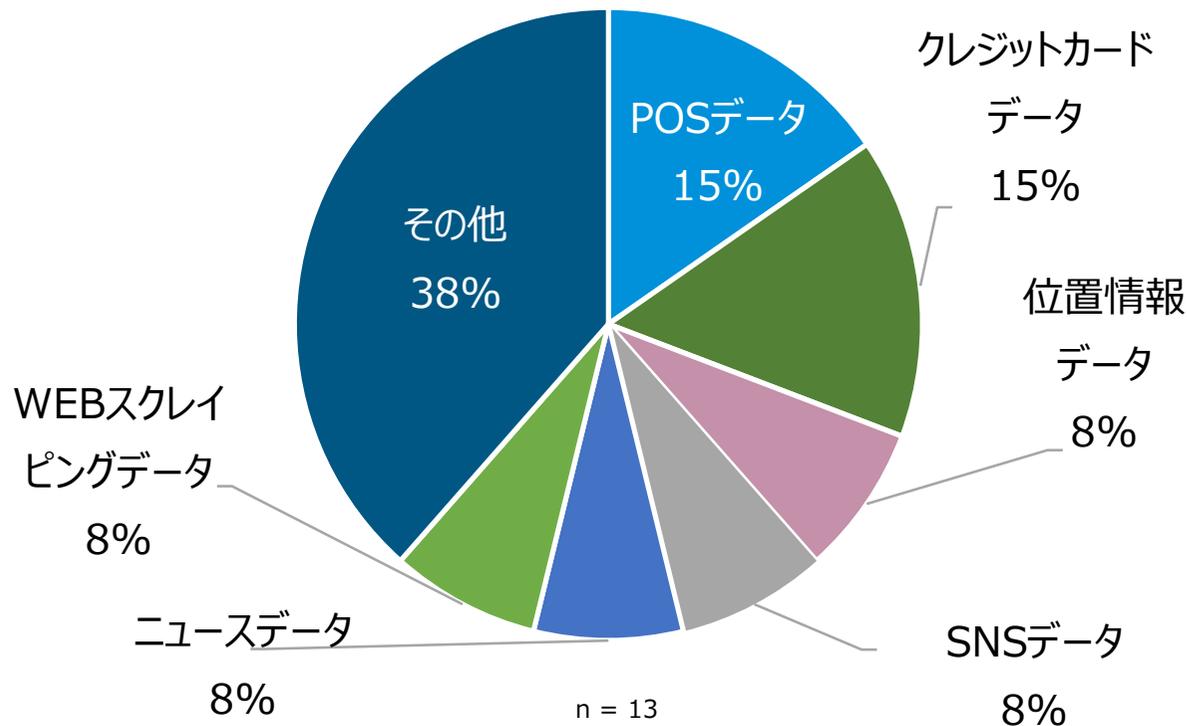
- ・資産価格データ
- ・口コミデータ
- ・市場関連データ
- ・医療データ
- ・特許関連データ
- ・従業員エンゲージメントデータ
- ・ESG関連データ
- ・アクセスログデータ
- ・テレビ視聴率データ
- ・決算関連データ



# 利用頻度が最も高いデータ

■ 購入者を対象に、最も利用頻度が高いオルタナティブデータを確認したところ、「POSデータ」や「クレジットカードデータ」が多く挙げられた。また、「位置情報データ」「SNSデータ」「ニュースデータ」「ウェブスクレイピングデータ」などを挙げる先も見られた。

## ■ 最も利用頻度が高いオルタナティブデータは何か（購入者）



## ■ その他

- ・資産価格データ
- ・口コミデータ
- ・市場関連データ
- ・スマホアプリデータ
- ・地理データ

# WEBスクレイピングでの取得データ

購入者

データプロバイダー

データ分析受託

- WEBスクレイピングの活用者を対象に、取得しているデータと活用方法を確認したところ、「企業の開示情報」を含めた各種公開情報や「X（旧Twitter）データ」など、多様なデータがスクレイピングの対象となっていることが分かった。

## ■ WEBスクレイピングで取得しているデータと活用方法は何か（WEBスクレイピングの活用者）

企業開示情報。

クラウドファンディング等に関するデータ（資金の調達状況等の動き）。

求人広告に関するデータ。

X（旧Twitter）データ。

自社と他社のシェアオフィス出店状況。

企業の開示情報。

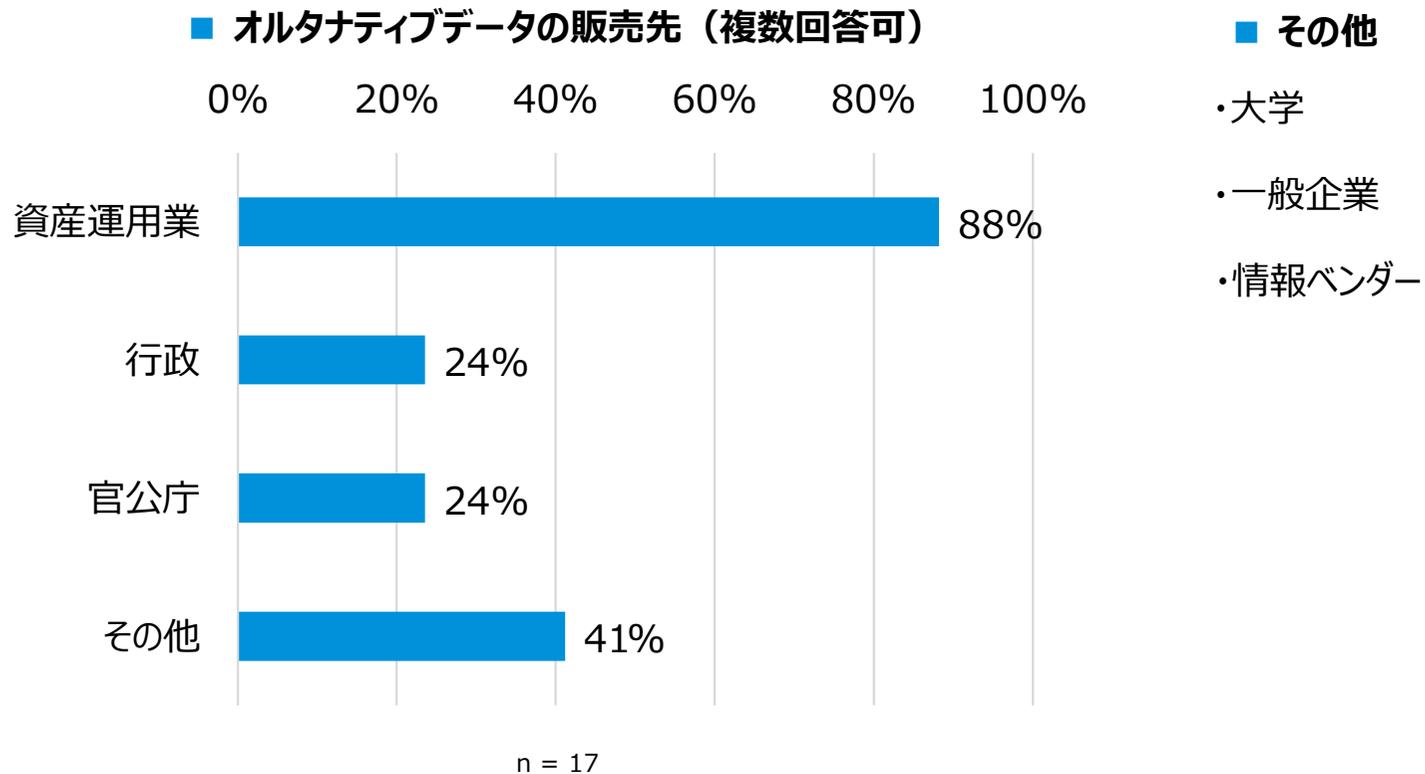
サイト上の各種発表情報。

AI・アルゴリズムを活用して、機械的データ収集し、クレンジング・名寄せ・名分けをして、データを提供したり、スコア化して提供する。

公開 Web サイトをスクレイピングし、サードパーティのデータプロバイダーを介してライセンスされたコンテンツを取得、自社データに活用。

# 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た。
- 最も多かったのは「資産運用業」向けで、「行政」や「官公庁」との回答もみられた。

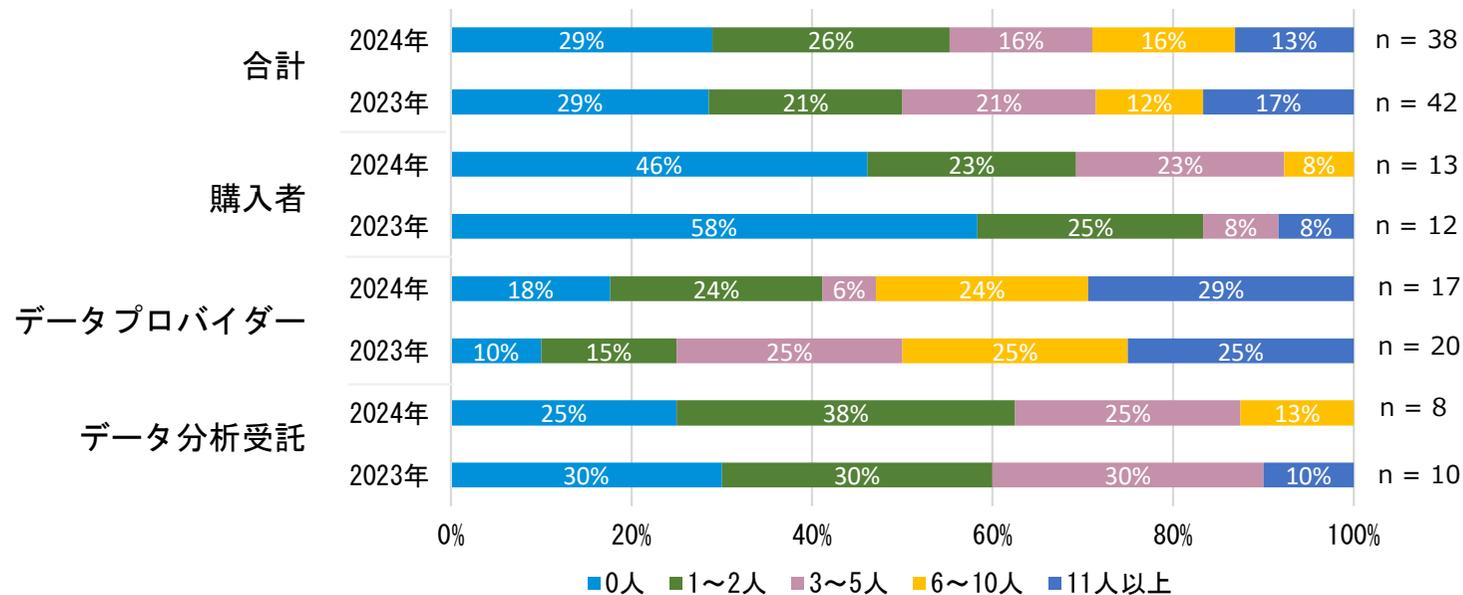


# オルタナティブデータ専門の人員

購入者      データプロバイダー  
データ分析受託

- オルタナティブデータを専業で扱う人員数を質問したところ、各主体で大きく傾向が異なった。
- データプロバイターやデータ分析受託では多くが専業の人員を有している一方、購入者は専業の人員が少ない「二極化」の傾向が示された。
- ただし、購入者については昨年調査と比較して、専業の人員を置く先が増加しており、体制の整備が進展していることが示されている。

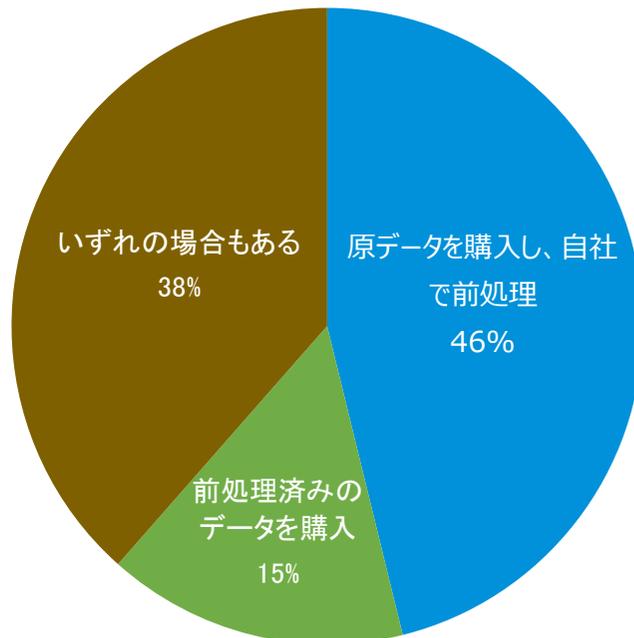
■ オルタナティブデータ専門の人員は何人いるか



# オルタナティブデータの前処理

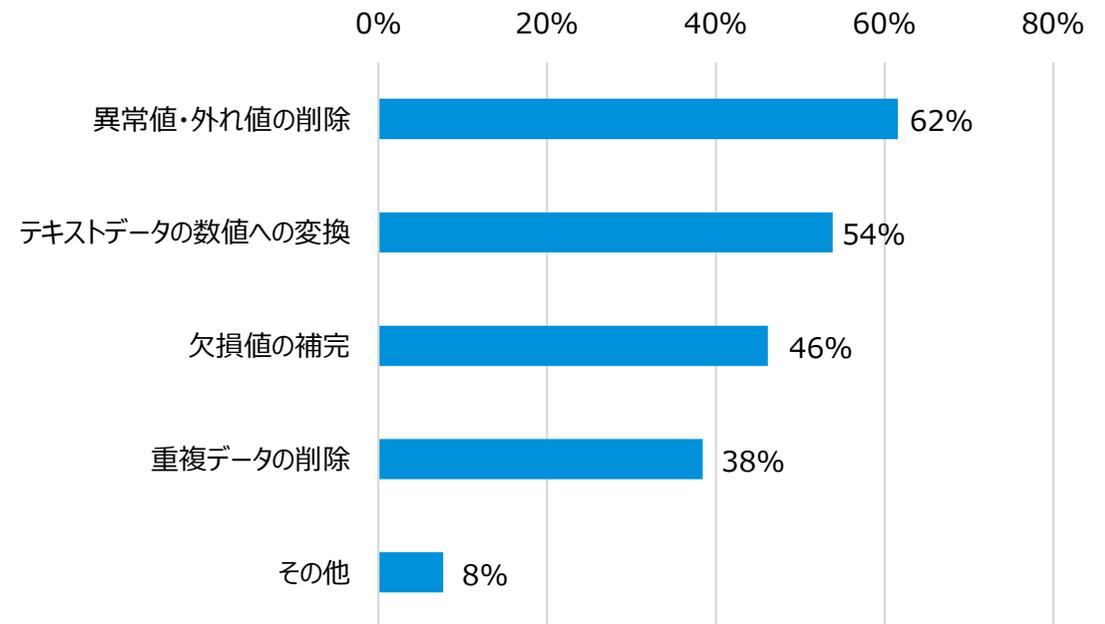
- 購入者を対象に、購入したオルタナティブデータの前処理について質問したところ、多くの回答者が自社で前処理を実施していることが示された。
- 前処理の内容としては、「異常値・外れ値の削除」が最も多く、「テキストデータの数値への変換」「欠損値の補完」などが続いた。

## ■ 購入したオルタナティブデータの前処理はどうしているか



n = 13

## ■ どの部分に前処理を行っているか（複数回答可）



n = 13

# 整備が必要な法律

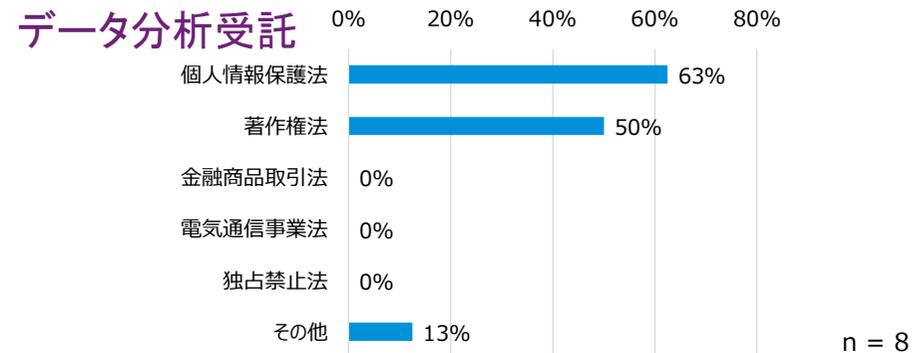
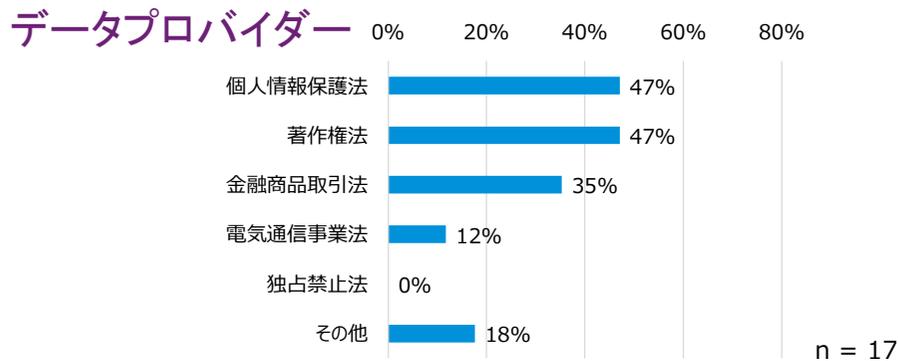
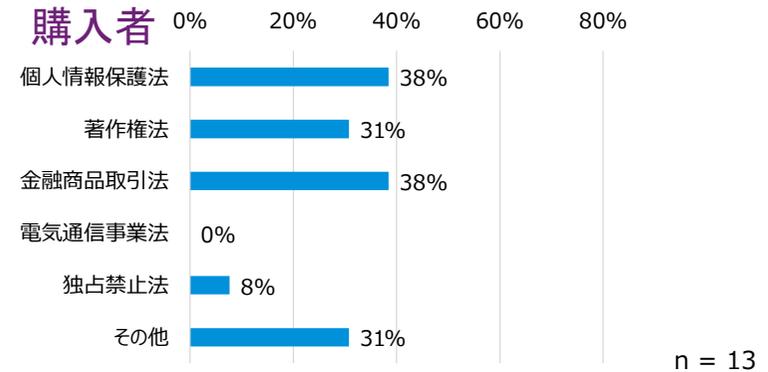
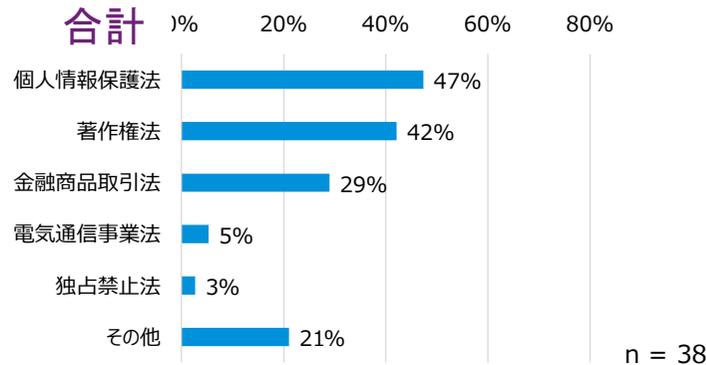
購入者

データプロバイダー

データ分析受託

- オルタナティブデータを利用するにあたって、整備が必要と思われる法律に関する質問では、「個人情報保護法」「著作権法」「金融商品取引法」が多く挙げられた。
- 主体別には、データプロバイダーやデータ分析受託は、特に個人情報保護法や著作権法に関する整備が必要であると考えていることが示された。

## ■ オルタナティブデータの利用にあたって整備が必要と思われる法律（複数回答可、最大3つまで）



# 整備が必要な法律（データ種類別）

購入者 データプロバイダー  
データ分析受託

- 取扱いしているデータの種類の（複数回答可）と、整備が必要と思われる法規制（複数回答可、最大3つまで）のクロス集計を行った。
- 個人情報保護法は、データの種類にかかわらず整備が必要であるとの回答が多かった。対して、著作権法ではPOSデータやニュースデータの利用ユーザーで法整備が必要であるとの回答が多い一方で、クレジットカードや位置情報では回答割合が低かった。
- 独占禁止法や電気通信事業法については、いずれのデータも回答割合は低くなっている。

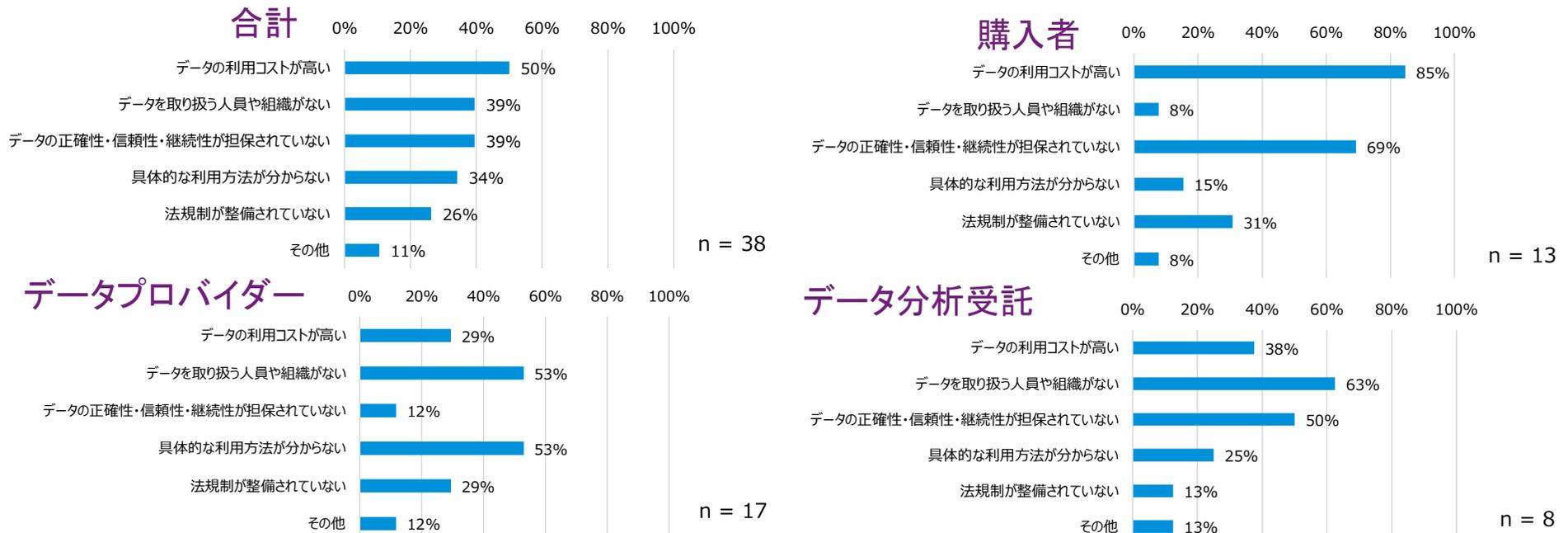
## ■ 取扱いしているデータの種類の整備が必要と思われる法律（複数選択可）

	個人情報保護法	著作権法	金融商品取引法	独占禁止法	電気通信事業法	n
POS	43%	57%	29%	14%	0%	7
クレジットカード	43%	14%	14%	0%	0%	7
位置情報	30%	10%	20%	0%	10%	10
衛星画像	33%	33%	33%	0%	0%	3
SNS	56%	33%	44%	0%	0%	9
WEBスクレイピング	56%	33%	44%	0%	0%	11
ニュース	43%	57%	43%	7%	0%	14

# オルタナティブデータの課題

- 課題について、全体では「利用コストが高い」、「オルタナティブデータを扱う人員がない」、「データの正確性・信頼性・継続性が担保されていない」が多く挙げられた。
- 購入者では、「利用コスト」が最多。データプロバイダーでは、「人員」「具体的な利用方法」が最多。データ分析受託では、「人員」が最多であるが、「正確性・信頼性・継続性の担保」「利用コスト」など多くの課題が挙げられた。

## ■ オルタナティブデータを利用するにあたっての課題（複数回答可、最大3つまで）





# 課題（購入者）

購入者

## ■ オルタナティブデータ利用の課題（購入者、自由記述）

データの蓄積が十分でないことで、統計の癖を把握することや季節調整をかけることなどが難しい。

正確性などが劣るため、主要なデータソースにしにくい。

オルタナティブデータの具体的なユースケースがはっきりしないまま、高い価格で販売されている印象がある。

入手したデータが特定の条件下のみのデータとなるため、汎用性は低い。間違った判断に至る可能性がある。

伝統的なデータに比べ費用が高い。

生成AIでの利用が禁止されており、プロジェクトごとに料金を払わなければならないデータがある。

社内で費用対効果を問われた時に、オルタナティブデータをお試しで利用してみることができないため。

500万円以上かかるデータが多く、リターン効果が十分に見込めないと導入は厳しい。

オルタナデータは新しいものが多いのでバックテストを行うのに十分な期間がない例が多い。

# 課題（データプロバイダー）

## ■ オルタナティブデータ利用の課題（データプロバイダー、自由記述）

従前の業務フローで想定されていないデータのため、顧客側で利用方法がイメージしづらい。

既存データとは異なる種類のデータは、個別の取り扱いが必要になる。

データを扱える人材の育成やデータそのものの普及には時間を要する。

データが高価であるため、価格以上のメリットを示すか、ディスカウントしなければ購入に繋がらない場合が多い。

多くの日本企業（特に一般事業会社）にはオルタナティブデータを利用することのできる人材が不足している。

もっと気軽に取り扱いができるようにする必要がある。

特に国内企業のオルタナティブデータ理解がまだまだ進んでいない。

分析を行う体制や予算が十分でないという感触がある。

データが高度であるが故に使い方がわからないことが多い。



# 課題（データ分析受託）

データ分析受託

## ■ オルタナティブデータ利用の課題（データ分析受託、自由記述）

- |  |
|--|
| 非構造化データの取扱いについての技術体系が整備されていないので、各データごとにかなりの工夫が必要になる。 |
| データ収集の前提がわからない場合、データの信頼性や正確性についても疑問が残る。              |
| データを取り扱う人員を採用することが困難。                                |
| データの正確性をProofするためのPoCが必要となり工数がかさむ。                   |
| 次から次へとデータのトレンドが変化していくため、その一つ一つを追うことはビジネス面で極めて負担が大きい。 |
| データの加工・分析のスキルを持つ人材が不足している。                           |

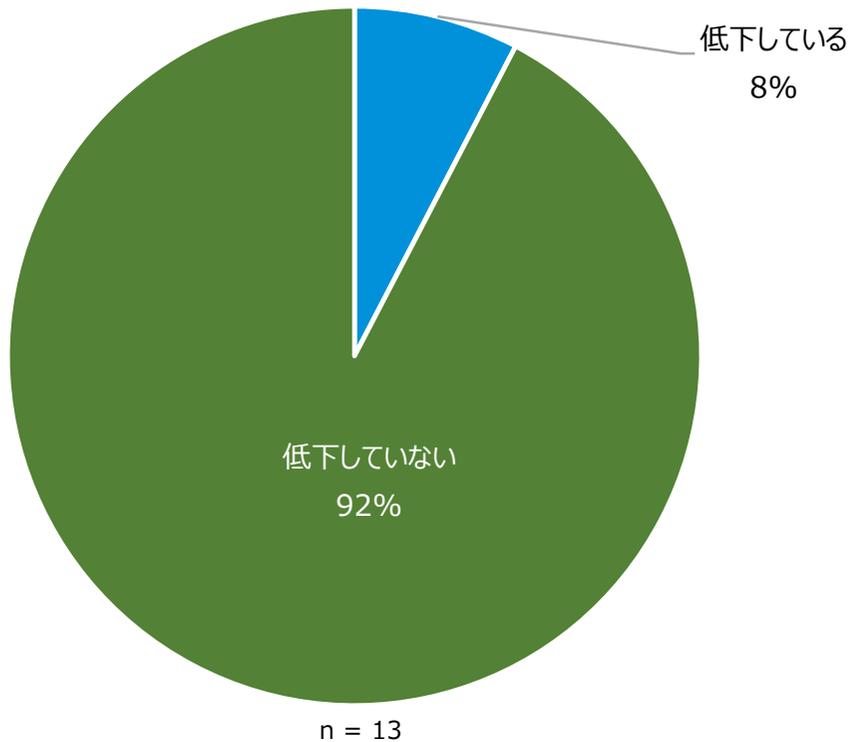
# オルタナティブデータに対するニーズ

購入者

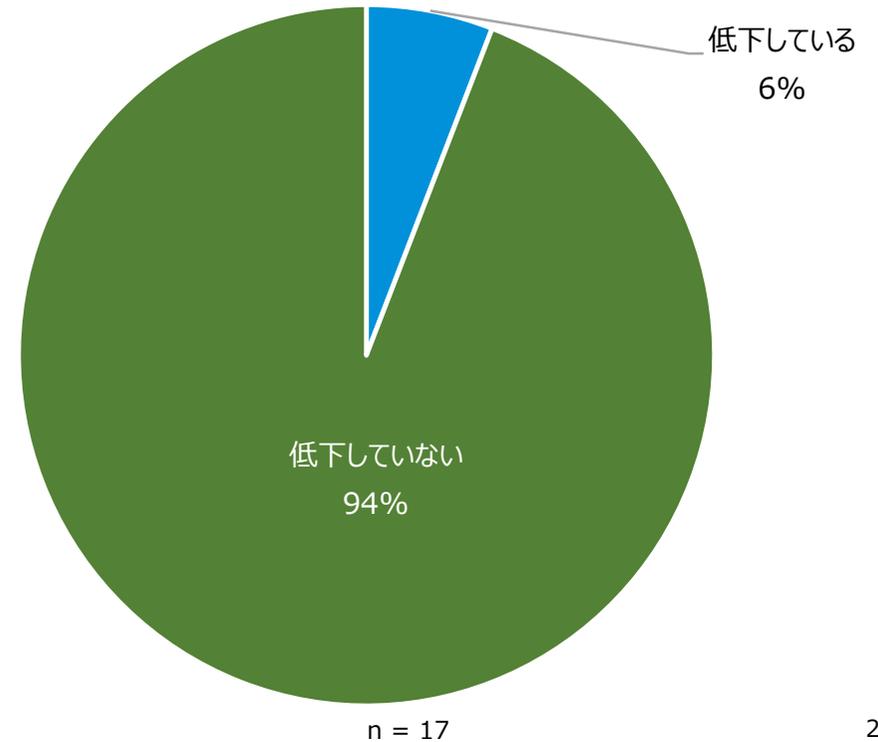
データプロバイダー

■ オルタナティブデータに対するニーズについて、購入者に利用ニーズを、データプロバイダーに顧客ニーズを質問したところ、いずれも「低下していない」との回答が多数を占めた。コロナ特需が終わった後においても、オルタナティブデータに対するニーズは引き続き強いことが示されている。

## ■ オルタナティブデータの利用ニーズは低下しているか



## ■ オルタナティブデータの顧客ニーズは低下しているか





# ニーズが低下していない（している）分野

購入者

データプロバイダー

## ■ ニーズが低下していない分野

資産運用業・投資銀行・保険業・事業会社
消費者行動・マーケティング、サステナビリティ関連、ヘルスケア
新聞、SNS等のテキストデータ
賃金や雇用に関する分野
生成AIでの活用やマクロ市場での活用が期待できるデータ
マーケティング領域や保険（未病）領域
金融商品取引データ
非財務情報
営業分野
保険料率
企業情報

## ■ ニーズが低下している理由・分野

コストの高いデータ
資産運用分野

# 今後の見通し

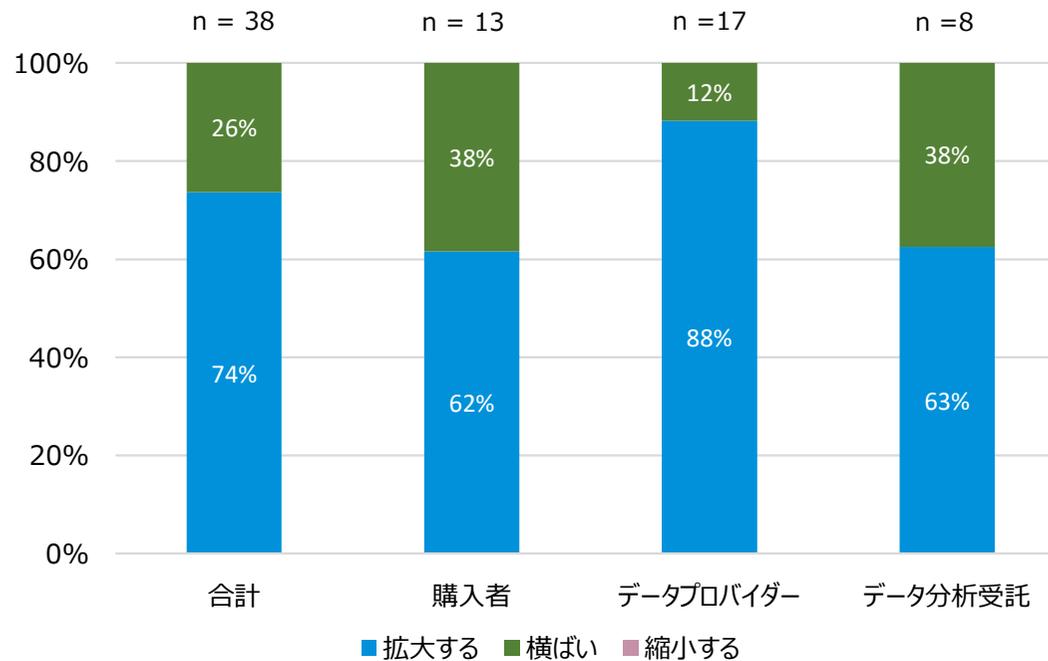
購入者

データプロバイダー

データ分析受託

- 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市場の動向については、全体、各主体共に、「市場は拡大する」が多数を占めた。

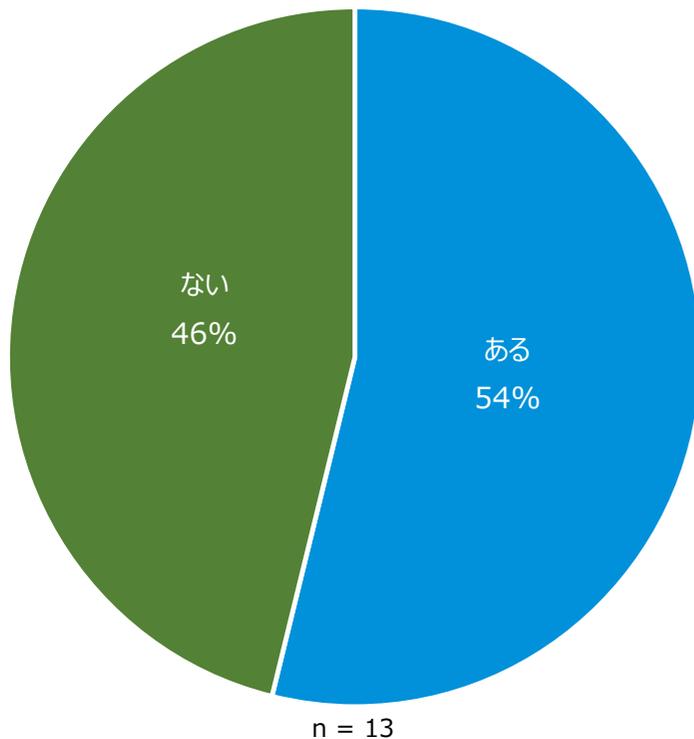
## ■ 今後（今年度含む向こう3年程度）、オルタナティブデータ市場は拡大すると思うか



# オルタナティブデータのAIへの活用

- オルタナティブデータのAIへの活用について、「活用経験あり」が54%と、「活用経験なし」をわずかに上回った。具体的には、「投資判断プロセスの効率化」「テキストデータの翻訳」「クオンツファンドへの活用」といった声が聞かれた。

## ■ オルタナティブデータを用いてAIをビジネスやリサーチ等に活用したことはあるか



## ■ 具体的な活用方法や利用データは何か

- ・投資判断やそのプロセスの効率化・高度化
- ・テキストデータの翻訳・要約
- ・クオンツファンドへの活用
- ・消費者行動の分析
- ・モデルの開発

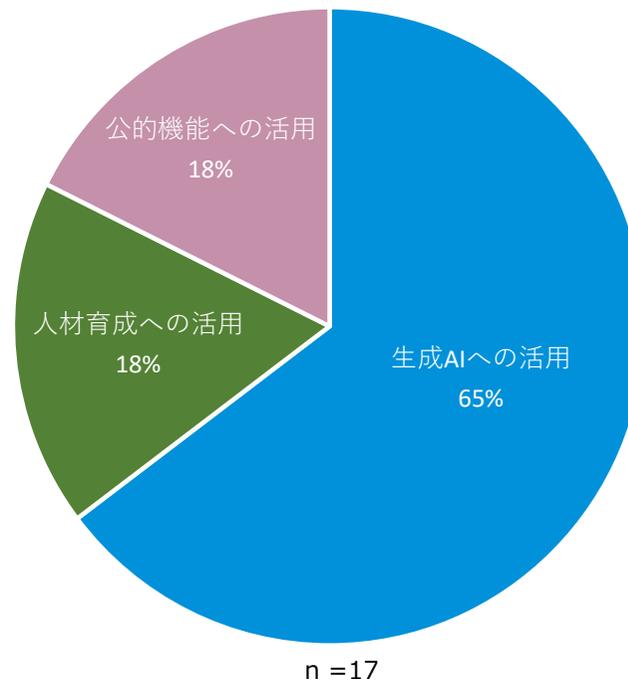
## ■ 活用していない理由は何か

- ・AIは活用しているが、オルタナティブデータとの結びつきはない
- ・具体的な活用事例がイメージできない

# 今後のユースケース

- データプロバイダーを対象に、今後伸びていくと考えられるオルタナティブデータのユースケースを聞いたところ、「生成AIへの活用」が多数を占めた。この他、「人材育成への活用」や「公的機能への活用」との回答も挙げられた。

- オルタナティブデータのユースケースとして、今後どのようなものが伸びていくと考えられるか

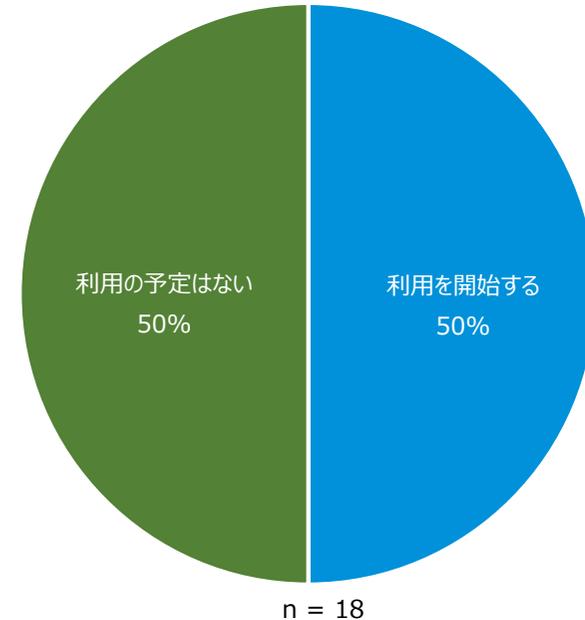
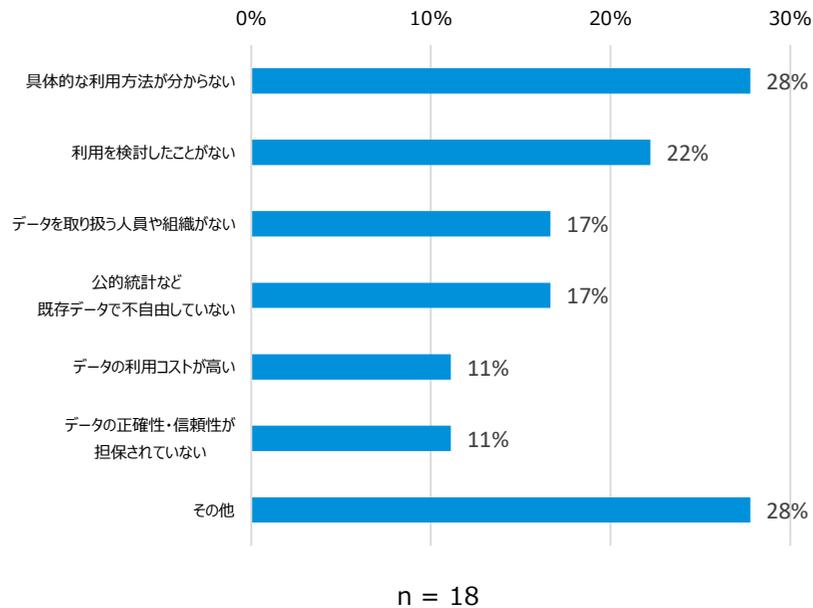


# オルタナティブデータを利用していない理由

未経験

- 「オルタナティブデータの利用経験がない」との回答者が利用していない理由については、「具体的な利用方法が分からない」、「利用を検討したことがない」、「データを取り扱う人員や組織がない」、「公的統計など既存データで不自由していない」などが多く挙げられた。
- 自由記述では、「自社でデータを取得・蓄積したいが組織が整備されていない」、「業務で活用できる事例などが把握できていない」などの記述があった。
- もっとも、未利用先でも、半数が今後「利用を開始したい」との意向を持っている。

## ■ オルタナティブデータを利用しない理由（複数回答可、最大3つまで） ■ 今後（今年度含む向こう3年間）のオルタナティブデータ利用の方針



## おわりに

- 本調査を通じて、「オルタナティブデータの利用が拡大している」ということが確認された。
- オルタナティブデータ市場は伸びが鈍化しつつも拡大を続けており、オルタナティブデータへのニーズも縮小する様子はみられていない。今回調査では、購入者における専門人員の拡充やAIの活用といったデータの利活用面での広がりも確認された。
- オルタナティブデータの利用にあたっては、下記のような課題が改めて明らかになった。
  - ①データの利用コストやインフラ整備コストが高い
  - ②データを扱う人員や組織がない
  - ③データの正確性・信頼性・継続性が担保されない

回答先であるオルタナティブデータ推進協議会の会員からの協力が得られれば本調査は継続して行い、オルタナティブデータの動向を定点調査していきたい。今後の調査に向けて、ご感想やご意見等をいただくと幸いです。



**JADAA**

オルタナティブデータ推進協議会

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