AIMA

SS&C

CASTING THE NET

How Hedge Funds are Using Alternative Data

CONTENTS

Acknowledgements

We are very grateful to the following AIMA research committee members for their dedication towards the creation of this document:

Carol Ward
(Chairperson)
COO Man GLG
Man Group

Tom Kehoe
Managing Director,
Global Head of Research
and Communications
The Alternative Investment
Management Association

Matthew Newbon
Chief Operating Officer
Independent View BV

Joanne Matthews
Senior Vice President
Two Sigma Investments LP

Hardik Shah
Business System Consultant
CIBC

Michael Peltz
Global Head of Content
WorldQuant, LLC

Tess Shih
Executive Director
Capital Fund Management
International Inc

Waheed Aslam
Head of Marketing &
Business Development
– Asset Management
& Investments Funds
Simmons & Simmons LLP

To the following third-parties for their valuable insights:

Gene Getman
Client Portfolio Manager
LOIM, 1798 Alternatives

Ronan Crosson
Director, Data Strategy
& Analytics
Eagle Alpha Ltd.

And to the following members of the SS&C team for their valued expertise:

Kelly Ramsey Gooch
Director, Relationship
Management
SS&C Technologies

Michael Megaw
Managing Director
Regulatory Analytics and Data
SS&C Technologies

Alastair Hewitt
Director or CORE-SightLine
SS&C Technologies

Co-authors:

Lyndsay Noble
Director of Data Science
SS&C Technologies

Anton F. Balint
Associate
The Alternative Investment
Management Association

FOREWORD

The global economy and financial markets are always changing. With them, the information that hedge fund managers can gain from analysing the world around them also evolves. Consequently, the tools needed to extract data from such information need to adapt – successful investing, irrespective of what strategy or style one employs, depends to a good extent on gaining and maintaining a legitimate information edge over the rest of the market. To put it differently, for hedge fund managers to meet the investment needs of their clients, they need to have a greater understanding of how the world functions than their competitors.

While traditional sources of economic and financial knowledge, such as textbooks, industry literature and established data bases are excellent in providing a level-playing field for hedge fund managers, going above and beyond these commonly used sources is crucial for managers to remain innovative and therefore, to stay competitive. In doing so, more and more alternative investment funds are adopting a 'quantamental' approach, a blend of fundamental investing combined with a more quantitative approach. Central to this new way of thinking is the emergence of alternative data.

As a concept, alternative data is not new: for thousands of years market-savvy business people have tried to understand their trading environment by looking at the world around them through different lenses and, from their observations, extracting data that, although not conventional, helped them to navigate the market successfully.

However, in recent years, enabled by the technological advancements across a number of industries, accessibility to alternative data sets has improved tremendously: with a growing number of alternative data providers, hedge fund managers now have access to a large number of non-traditional data sources, such as satellite imagery, social-media trends and weather patterns.

The aim of this publication is to offer an in-depth analysis of this topic, as well as to invite all stakeholders interested in how alternative data is being used by the hedge fund industry to further discuss its broader adoption. In it, you will discover how widely adopted alternative data is within the hedge fund industry, what are the main uses that managers are employing alternative data for, the opportunities and challenges that these data sets present and what the future holds for alternative data within the hedge fund sector.

We would like to thank AIMA's research committee for their valuable input and for taking the time to discuss these findings. We would also like to thank the various asset managers for their generosity in contributing the several testimonials included throughout this paper, and to Eagle Alpha for their insight. Finally, thank you for taking the time to read this paper. We would love to hear your thoughts.

Jack Inglis

Chief Executive Officer,
Alternative Investment
Management Association

Michael Megaw

Managing Director
Regulatory
Analytics and Data
SS&C Technologies

INTRODUCTION

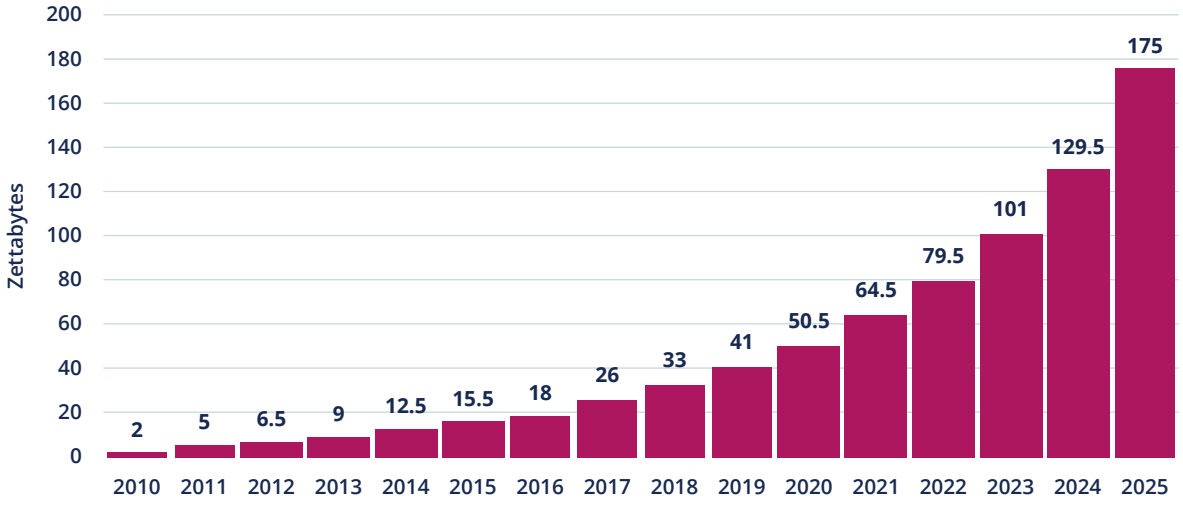
Change is the only constant in the known universe. Nowhere else is this truer than in the world of asset management, especially in alternative investments. Within this space, hedge funds continue to adapt to an evolving landscape of challenges and opportunities. Among other things, this includes adopting novel technologies for managing risks, researching investment ideas and, ultimately, generating alpha.

The Alternative Investment Management Association (AIMA), being well-positioned at the heart of the industry, has been witnessing this transformation first-hand. Moreover, AIMA has been assessing the impact of technology on hedge fund managers and their clients through a series of in-depth research papers, including the landmark publication "Perspectives". This paper continues AIMA's work in this space and, in collaboration with global fund service provider SS&C, it looks at how hedge fund managers are using alternative data.

Over the past thirty years, the number of alternative data providers has grown from about 20 in 1990 to just over 400 in 2018[1]. This expansion has been driven by a number of factors, including the increase in the amount of data itself, the proliferation of new data sources, improvements in computational power and advancements in data science. As International Data Corporation (IDC) put it in a November 2018 study, "mankind is on a quest to digitize the world". Indeed, IDC estimates that the size of the global datasphere in 2025 will reach 175 zettabytes[2], an increase of about 4.3 times from 2019. Alternative data is, obviously, part of this huge information universe.

1 See here: https://alternativedata.org/stats/
2 A zettabyte is equal to 10 at the power of 12 gigabytes or 10 at the power of 9 terabytes. To contextualise what 175 zettabytes look like, the largest hard drive in the world is 15 terabytes in capacity – in the absence of the cloud technology, we would need 11.7 billion of the world's largest hard drives to contain all that data.

SIZE OF GLOBAL DATASPHERE



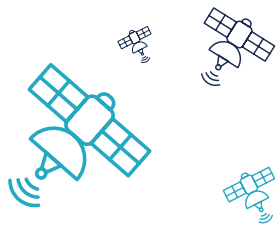
Source: IDC

Defining alternative (alt) data is not easy. Indeed, the lack of a universally accepted definition can be a barrier for both hedge fund managers and regulators when looking to assess the risks and rewards presented by alternative data.

In the simplest sense, alternative data can be regarded as everything that doesn't fall within the realm of traditional financial and economic data. However, this is a broad and non-practical way of looking at alternative data. Therefore, in deciding on a definition we must focus primarily on its practical application.

As a concept, alternative data is not new. Indeed, it goes as far back as ancient Babylon when merchants used measurements of the Euphrates' depth and flow to inform their decisions in trading various commodities, as they realised that these variables were correlated with market supply[3]. However, what is new in recent years is the increasing level of accessibility to this type of data. With a growing number of alternative data providers, hedge funds have access to a myriad of data sources, such as satellite imagery, social-media trends and consumers' shopping behaviour.

In the simplest sense, alternative data can be regarded as everything that doesn't fall within the realm of traditional financial and economic data.



3 See here: https://caia.org/sites/default/files/014-031_monk_jfds.pdf

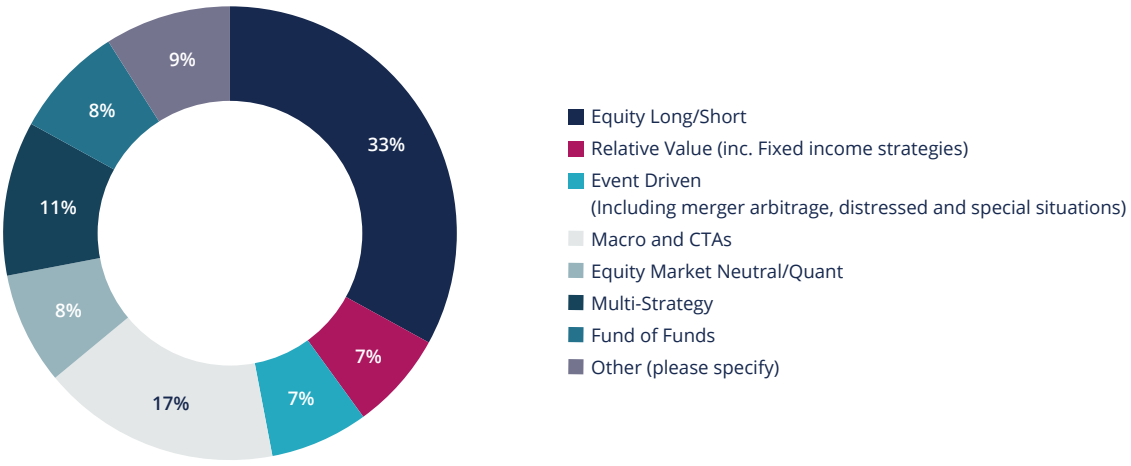
Considering how fast technology is moving, those that fail to adapt risk losing a potential advantage in the competitive race to deliver the rarest of returns – alpha. Consequently, most managers are in the process of updating their investment processes and business models in order to accommodate the growing amount of alternative data.

We hope you find this publication insightful and useful and we invite other stakeholders to join the conversation around alternative data within the hedge fund industry.

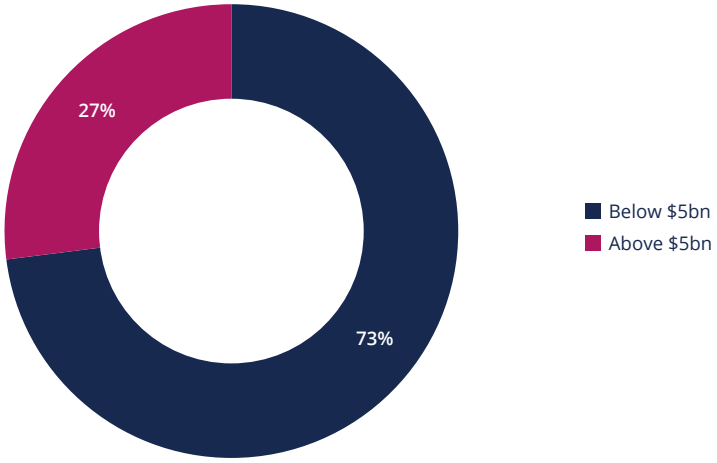
METHODOLOGY

In order to collect the necessary data, we ran a survey to which 100 hedge fund managers responded, managing a total of about $720bn in assets. Additionally, we have collected insights from conversations with managers and alternative data providers.

HEDGE FUND STRATEGIES OF RESPONDENTS (TOTAL POPULATION)

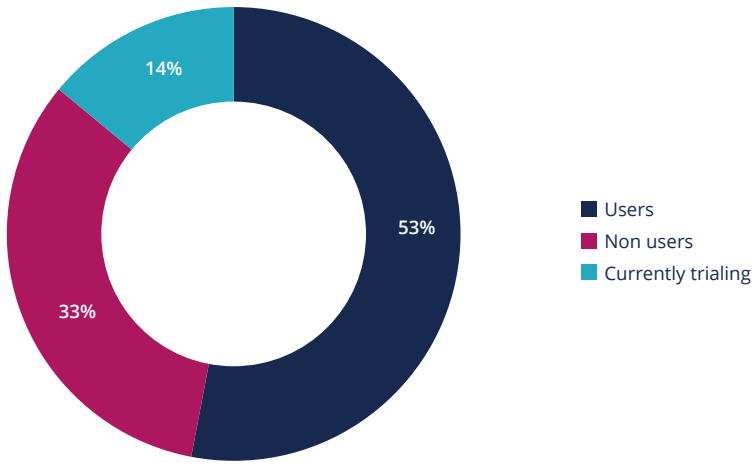


ASSETS UNDER MANAGEMENT OF RESPONDENTS (TOTAL POPULATION)



To better understand if and how the hedge fund industry has been adopting alt data, we split the survey data first into users (53% of all respondents) and non-users (47% of all respondents). Over half of our respondents are users of alternative data. In our classification of users and non-users, the managers which are only trialling alt data options are classified as non-users. These represent 14% of the total population. Consequently, the 53% figure represents only managers that are using alternative data – these are either heavy users (23% of total population, or 45% of the users' sample) or light users (30% of the total population, or 55% of the users' sample). When we aggregate the data, we get the overview below.

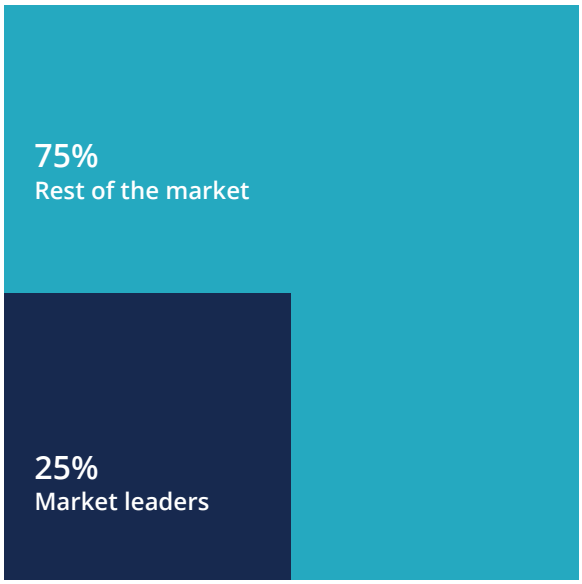
USERS AND NON USERS (TOTAL POPULATION)



We further split the users into "market leaders" – defined as those managers that have been using alternative data for more than five years – and the "rest of the market", which includes all those respondents that have been using alternative data for less than half a decade. Not surprisingly, 92% of the market leaders are also heavy users of alt data.

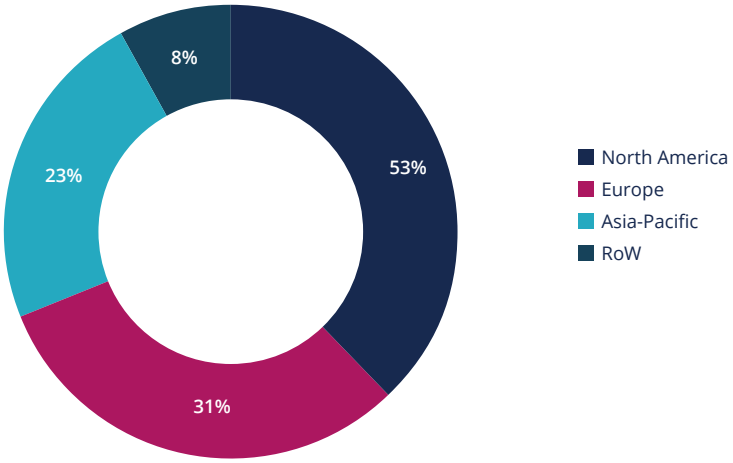
Out of the total population of responses, only 13% are classified as market leaders (or just about 25% of the users' sample). We expected this: The usage of alternative data follows the same adoption curve as any other novel concept or technology. However, as mentioned in the introduction, this type of data is not strictly new. What is new is the codification and quantification of data from unconventional sources of information on a large scale and its availability to hedge funds.

USERS SPLIT BY "MARKET LEADERS" AND THE "REST OF THE MARKET"

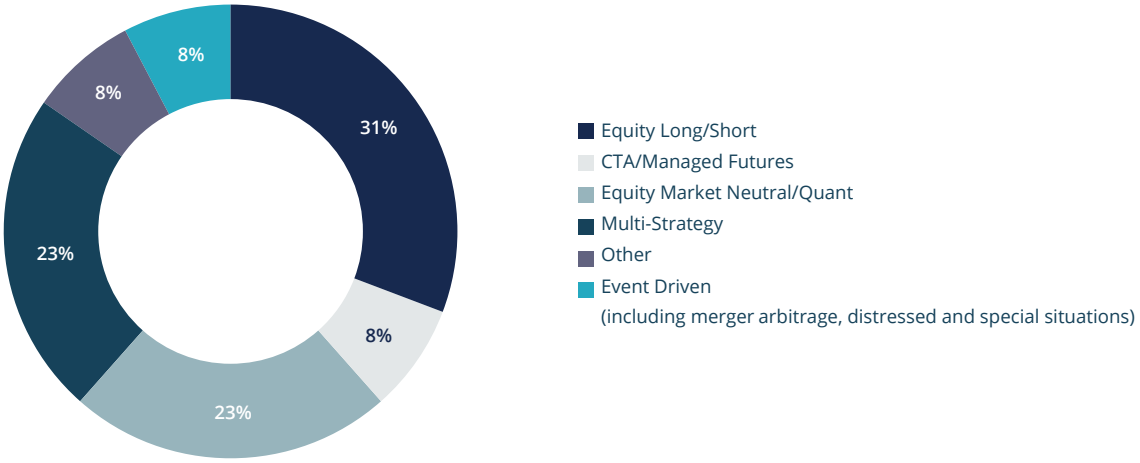


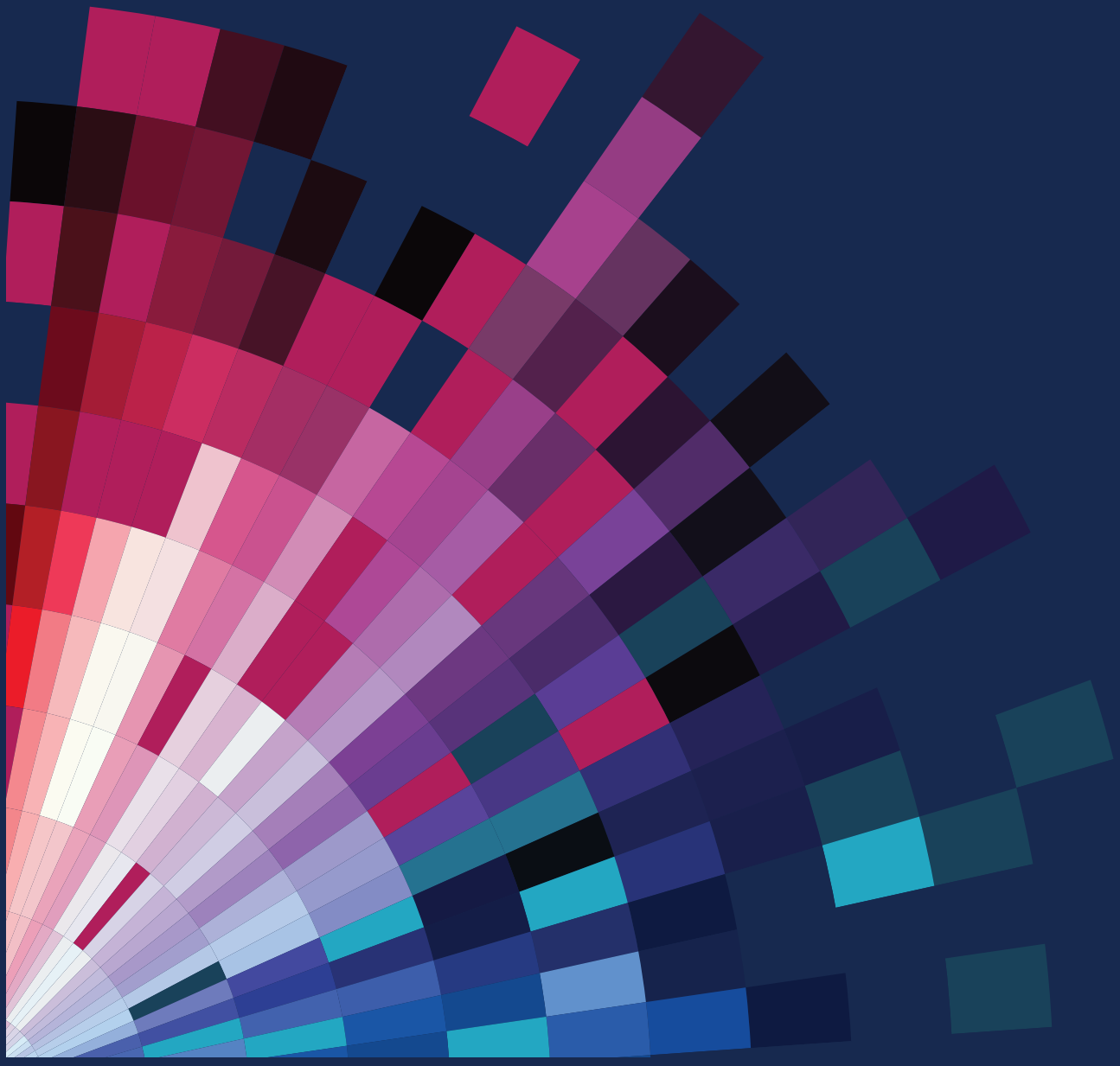
Interestingly, when we look at the assets under management (AUM) of market leaders, the split is 46%-54% between managers running more than $5bn and those that manage less than that. This, however, is likely to be the result of our sample size, rather than an accurate reflection of the broader market – in practice, common sense dictates that larger players will have more resources available to dedicate to alt data. Nevertheless, it illustrates that there are smaller managers which are nimble enough to remain innovative in their hunt for alpha.

MARKET LEADERS BY THEIR LOCATION



MARKET LEADERS BY STRATEGY





CHAPTER 1

DEFINING ALTERNATIVE DATA

The role of a definition is to provide clarity when using a concept, a good or a process. In the case of alternative data, one way of achieving such clarity is by looking at its practical application.

In a recent paper, consulting firm Deloitte explains that "alternative data augments traditional asset allocation models by enabling additional insights into the investment opportunity universe. It also allows for an improved risk management process"[4]. Meanwhile, Eagle Alpha, an alternative data provider, states that "alternative data is not traditional data"[5] – this is closer to the definition offered by technology company Oracle as "any information that is non-market data"[6], as well as by the website Alternativedata.org: "data used by investors to evaluate a company or investment that is not within their traditional data sources."[7]

However, questions remain as to what is traditional or market data? Can alternative data, if adopted by enough users, become "traditional"? In a sense, we all know what we mean by traditional data – the widely used economic and financial information provided by national governments, international institutions or companies which includes things such as employment figures, GDP numbers, accounting reports and so on. Essentially, traditional data can be thought of as the data that academic textbooks and asset management industry literature portray as the basis for economic activity and capital allocation decisions.

Alternative data, in a broad sense, can be thought of as "everything else" that asset allocators use to get information – in order to narrow this definition down to a level that is practical, we need to look at where the data comes from (its source), its structure, its distribution and its use.

When it comes to the data source, it is useful to think about whether the data in question comes from "conventional" or "unconventional" information. Something becomes conventional depending on whether it is accepted by a large enough group of people over a period of time that is long enough to become common knowledge, which is then passed as such from generation to generation through education and training.

4 See: https://www2.deloitte.com/content/dam/Deloitte/us/Documents/financial-services/us-fsi-dcfs-alternative-data-for-investment-decisions.pdf
5 See: Unlocking the potential of alternative data for corporates, Edition 1, July 2019
6 See: https://www.oracle.com/a/ocom/docs/industries/financial-services/searching-alpha-alternate-data-sets-wp.pdf
7 See: https://alternativedata.org/alternative-data/

Therefore, we can say that alt data comes from unconventional information which is not the common knowledge typically seen as key sources of economic and financial data.

Moreover, alternative data usually comes in a less structured format than traditional data, meaning that it doesn't come organised in a pre-defined way. Therefore, alt data is not easily searchable, as it includes images, videos, audio and social media posts. Consequently, in order to turn this kind of data into actionable insights, advanced algorithms, substantial computational power and enough cloud storage, as well as a "think different" mindset, are necessary – this makes the infrastructure needed to employ alternative data much more robust (and perhaps costly) than what is needed to work with traditional data.

Although alternative data is becoming more widespread across the asset management industry (indeed, it is already being used more widely within corporate sectors, such as technology), it is still a relatively niche source of information that is used by alternative asset managers. As such, this data tends not to be too widely distributed, unlike traditional data which has been widely made available through information sets managed either by public bodies or private actors.

There is a point where a particular data set, if it is too widely distributed for a sustained period of time, can become part of the common knowledge of the industry and therefore, through this process of commoditisation, alternative data becomes traditional data. As such, it is important for a data set to maintain its "niche" distribution aspect if it is to be considered alternative.

Finally, as we will see in the following chapter, alternative data is being used not just within the investment management process – be it at the research stage or when a capital allocation decision is being made – but also for business development purposes, including business risk management or improving internal operational processes.

Putting everything together: **alternative data comes from unconventional information, mostly in an unstructured form, is not broadly distributed within the industry and is being used to deliver both investment alpha and operational alpha.**

CHAPTER 2

THE OPPORTUNITIES

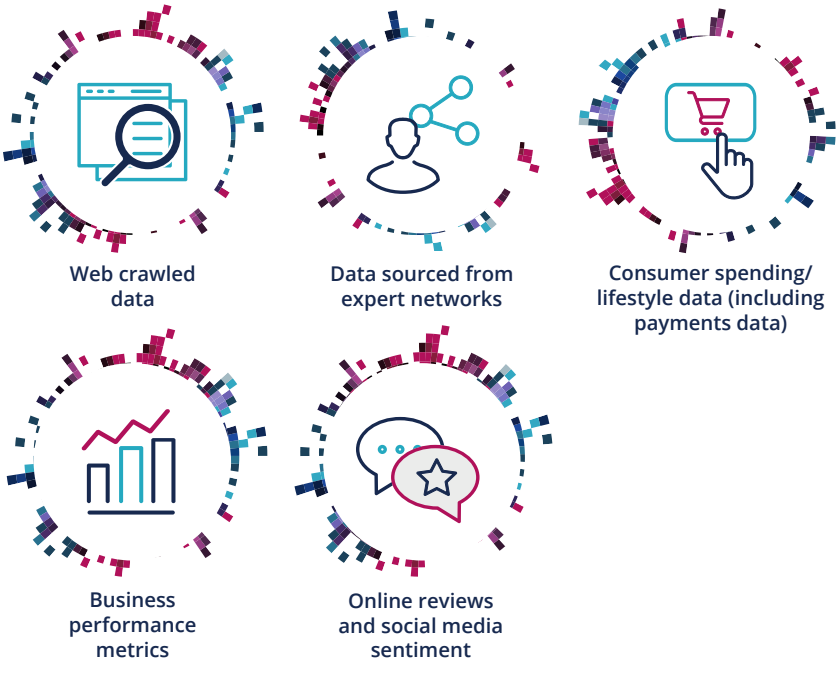
Alternative data sets used by hedge funds

When asked what are the most used alternative data sets, the market leaders responded with: data sourced from expert networks (this is bespoke research that may include data from unconventional sources), web crawled data, consumer spending/ lifestyle data and business performance metrics. The latter data set may come as a surprise to some readers as to why it is regarded as alternative.

However, from discussions with alternative data providers, we've learned that looking at traditional data in new ways – for example, developing unique valuation metrics which are not part of the conventional way of assessing asset risks and rewards – is regarded as alternative data. Although we would argue that this doesn't fully fit with our definition provided previously (as these business metrics are typically in a more structured format and come from conventional information), there are industry players (on the sides of both vendors and buyers) that regard such data sets as alternative, hence their inclusion here.

TOP 5 ALTERNATIVE DATA SETS USED BY "MARKET LEADERS"

From left to right the most used alternative data sets

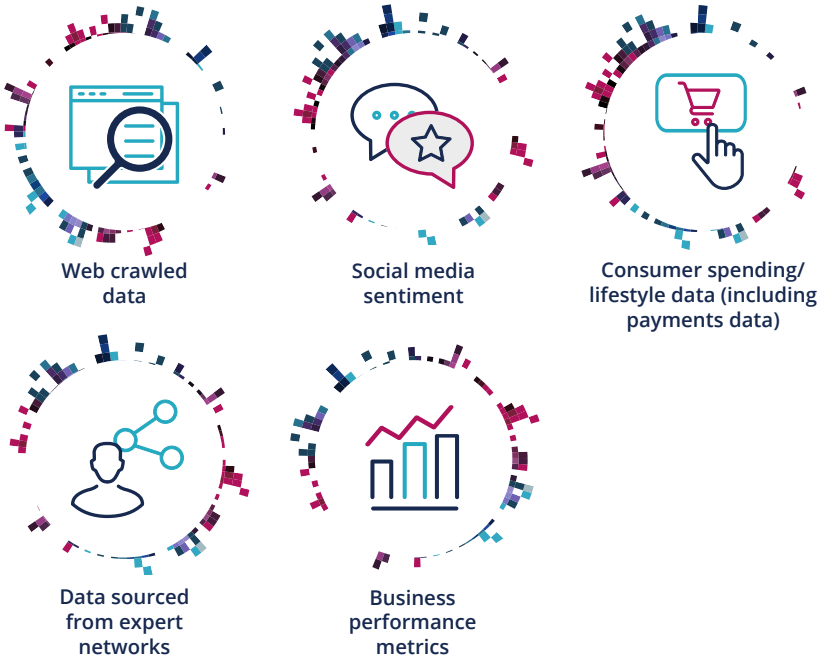


Analysing market leaders by strategy, we find out that long/short equity funds tend to predominantly use sentiment data, online reviews and payments data. Meanwhile, quantitative hedge funds prefer weather patterns, satellite imagery and logistics data, as well as web crawled data. Interestingly, the CTAs from our sample prefer only logistics data and consumer spending/lifestyle data. However, this analysis is not conclusive – rather, it merely illustrates just how diverse the alt data space is amongst hedge funds: there is no "one size fits all" solution.

The "rest of the market" tells a similar story – the top three most widely adopted alternative data sets are web crawled data, sentiment from social networks and consumer spending / lifestyle data. Although we haven't captured this in our survey, a key alt data set that sits between "sentiment from social networks" and "consumer spending" is "news analytics". Providers of market news, such as Dow Jones Newswires that have access to a large database of news stories from global media companies, as well as more local providers, are in good market position to offer hedge fund managers (especially those that run quant strategies) ultra-fast access to news stories – this can help them develop profitable trading models, optimise portfolio allocations and manage risk.

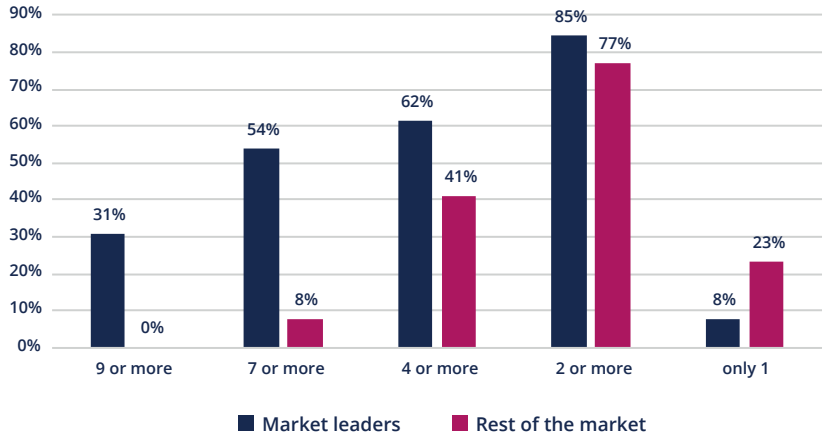
TOP 5 ALTERNATIVE DATA SETS USED BY THE "REST OF THE MARKET"

From left to right the most used alternative data sets



An interesting aspect to note is the breadth of data used by market leaders relative to the rest of the market, especially at the individual firm level. This means that a hedge fund classified as a market leader in the use of alternative data is likely using more such data sets than those that are later adopters. The chart below illustrates this clearly.

THE NUMBER OF ALTERNATIVE DATA SETS USED BY RESPONDENTS



With this in mind, we can now drill down on the opportunities that these alt data sets offer managers. As we mentioned above, alternative data can be used to deliver either (or both) financial and operational alpha. However, there are significant differences in how "market leaders" and "the rest of the market" are using alternative data.

MAIN USES OF ALTERNATIVE DATA

Nearly 70% of those classified as “market leaders” are using alt data in their search for investment alpha (versus just 44% of the “late adopters”). This suggests that, as hedge fund managers spend more time (and other resources, such as human and financial capital) on alternative data, cleansing it, structuring it and extracting valuable patterns, they progress towards employing this data in their search for investment alpha.

Market leaders are using more data sets than later adopters.

Another interesting insight derives from the bars in the right corner of the chart – some hedge fund managers are using alt data to improve risk management and compliance models. While the first thought would be around the risk management and compliance related to capital deployment, alternative data is also used to improve business operations (operational alpha). Our survey hasn’t captured this information directly. However, from our conversations with various industry players, we’ve learned that asset managers (some of which may be outside the alternative investment space) are using alt data as a tool to improve how their organisations function: for example, making business development decisions, such as the allocation of human resources, based on novel data analysis methods and (often bespoke) technology platforms that collect information about individual employees’ skills and output. This can help shape not just the productivity within the firm but also improve internal communication, due diligence processes and the overall culture.

Nevertheless, from our in-depth conversations with hedge fund managers and alt data providers, we learned that there are two large opportunities that alternative data presents to managers. Firstly, alt data represents an additional, valuable tool for helping hedge fund managers deliver outperformance. Secondly, it can also be used to enhance various risk management processes.

A tool for generating outperformance

Generating alpha –i.e. delivering a return that can be solely attributed to the hedge fund manager's skill rather than anything else – is the holy grail of any investment activity[8]. However, for alpha to be delivered, managers need to first identify and then exploit an investable edge – that is, they need to spot opportunities that other market participants either overlooked or completely missed and, importantly, allocate to them in a proportion that, at the overall portfolio level, more than makes up for the positions that do not go as well as initially thought. Alpha therefore is hard to come by.

How hedge fund managers go about using alternative data to help deliver outperformance depends primarily on their investment approach or style – discretionary or quantitative – and also on their strategies and underlining philosophical underpinnings to capital allocation. This is very well captured by ***Olga Kokareva*** from ***Quantstellation***, a multi-asset multi-horizon quantitative investment firm. In an **interview** provided by Eagle Alpha, Kokareva stated: ***"It's important to understand that usage of alternative data by fundamental hedge fund managers and by quantitative hedge funds are two very different processes. Fundamental hedge fund managers normally use alternative data to reinforce their investment thesis that they derived from their regular research process. For example, a manager can hold a long position in a retailer and they are thinking about closing it, but they are not sure. So, instead of waiting for the next quarterly report they can start looking at foot traffic data or credit card data. If the sales numbers are indeed going down, they might close this position earlier."***

Indeed, discretionary asset allocation rests on backing an informed belief either to exploit an opportunity or to avoid a risk. Therefore, for an asset manager that pursues this approach, the use of alternative data doesn't constitute the sole basis on which an investment decision is made or a trade placed; rather, it is another tool managers use to further enhance their judgment process about the world around. It is more difficult for discretionary managers to find out from where exactly each outperformance basis point came as the nature of belief is not quantifiable. In contrast, quantitative managers can be more specific as to how they employ alternative data in this regard. Kokareva continues: ***"Quantitative hedge funds derive their investment thesis by analyzing data, and they've been doing it for years. So that was a natural extension of the quantitative investment process – to start analyzing alternative data sets in exact same manner that they've been analyzing traditional financial data."***

In this survey, we didn't capture which respondents are discretionary and which are quantitative managers. However, we asked them to provide for what purposes they use alternative data and "helping to generate outperformance" and "researching investment opportunities" were among the top reasons (as the chart above also illustrates).

8 See: https://www.aima.org/educate/aima-research/perspectives-research.html

What stood out from the data, however, is that out of the total number of alternative data users, 10% of those are using "weather patterns" in their pursuit for outperformance. This sample is too small to draw any meaningful conclusions regarding the main strategies that are indeed extracting insights from this type of alternative data. However, as climate risk is likely to continue to increase[9], climate-related data will become more relevant for how hedge fund managers allocate capital. Similarly, ESG-related data (such as the ESG scores provided by Refinitiv[10]) – which sits outside traditional economic and financial information and therefore, can be regarded as a type of alternative data – will likely become a more substantial component in hedge fund managers' pursuit of alpha.

The process of delivering outperformance is complex, entailing various steps from conducting preliminary research on an investment opportunity, deciding on how much capital to allocate and, finally, executing the trade. Alternative data has the potential to enhance this entire pipeline of producing attractive returns. For example, employing alt data at the research stage can enable hedge fund managers to see the world from a different perspective: take the prospect of investing in the stock of a car manufacturer – from traditional data sources (like earnings calls, company reports and management presentations), information such as revenue streams, profit margins, R&D efforts and expansion plans can be extracted.

A great deal about the ecosystem within which the business operates remains hidden: Online reviews can reveal clients' perception of the company's brand, web crawled data can highlight new technologies that can impact the industry and which managers may want to explore as separate investments, and logistics data can offer perspectives on geographies in which the car manufacturer can export its product with more ease. These are all important information related to this investment opportunity, which traditional data sources do not capture.
As ***Mani Mahjouri, CEO of Blueshift Asset Management*** stated in an **interview** provided by Eagle Alpha, ***"It's all about building the model that describes how the world works."***

As we shall see later in this paper, however, extracting useful, actionable insights from alternative data sets – be it in the context of generating alpha or for other uses – requires hedge fund managers to have in place the appropriate infrastructure of technology and human talent. Assuming that such infrastructure is in place, we would like to list the main practical applications of alternative data in the search for alpha. This list is based on our desk research and includes: dynamic stock screening, targeted security baskets, precise sector allocation, portfolio optimisation, strategy backtesting and market timing.

9 See: https://www.wsj.com/articles/for-the-economy-climate-risks-are-no-longer-theoretical-11579174209?mod=hp_lead_pos11
10 See: https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/esg-scores-methodology.pdf

Discretionary Data (Ed Cole, Man GLG)

For a traditional discretionary stock picker, information on how a particular company is doing usually comes on a quarterly basis with the release of financial results. Of course, there are other data points – how competitors or suppliers are doing – but the most relevant data on actual execution at a company level comes with low frequency. So, how does a discretionary portfolio gain an edge?

In our view, use of alternative data – from app downloads and website traffic to social and news media – can potentially help a discretionary portfolio manager know about company execution sooner than its management, and certainly sooner than other investors who depend on brokers as conduits for information.

Let's focus on the luxury goods sector.

Luxury goods is almost uniquely a China-centric growth story. However, the analysts that cover the luxury goods companies – and the asset owners who own their shares – are almost entirely based in the EU and US. This discrepancy can create scope for information asymmetries. We can exploit these information asymmetries using natural language processing (NLP), which can help us systematically quantify consumer sentiment towards luxury goods brands.

Let's look at how NLP can be used in an example relating to Versace. In a nutshell, the process works as follows:

- The programme runs an online search for news mentions generated about Versace in China;
- This is translated into English using an online translator (note that the translation doesn't have to be in 'proper' English as we are looking for sentiment);
- NLP dictionaries analyse the text and rate in terms of polarity (i.e. positive, negative or neutral) and intensity (the level of positive or negative mentions).

The timeline with the Versace saga was as follows:

- 24 June, 2019: Versace is endorsed by popular actress Yang Mi; headlines generated yielded a positive NLP score of 0.42;
- Early July: A Versace t-shirt implying that Hong Kong and Macau are sovereign territories hits the shops;
- 24 July, 2019: Versace says it had stopped selling the t-shirts and destroyed them;
- 11 August, 2019: Yang Mi cuts ties with the brand; headline generates a negative NLP score of 0.74.

So how can discretionary managers apply NLP in their day-to-day stock picking? We believe this sentiment analysis can be used as part of a mosaic on information:

- Resilient sentiment versus weak stock price means contrarian long opportunity;
- Deteriorating sentiment in context of bullish consensus expectations means EPS downgrade risk;
- Sentiment at odds with our priors;
- Pair trade opportunities.

Consequently, alternative data can help discretionary managers in the long gaps between reporting periods. Used well, alternative data can potentially help discretionary managers stay a step ahead of management – and a step ahead of competition.

"The process of delivering outperformance is complex, entailing various steps from conducting preliminary research on an investment opportunity, deciding on how much capital to allocate and, finally, executing the trade. Alternative data has the potential to enhance this entire pipeline of producing attractive returns."

A risk management tool

Alternative data can also help hedge fund managers to manage risk – both investment risk and operational risk. Although only 10% of the users of alt data are seeing this as an opportunity, there are a growing number of industry and academic studies which demonstrate that alternative data can be used successfully as a risk management tool[11].

From a capital allocation perspective, managing investment risk efficiently is something that hedge fund managers are well-known for – the concept of risk-adjusted returns (the Sharpe ratio) is based entirely on the ability of the manager to deliver both attractive returns and to manage the risk exposure. As data provider Factset recognises, "alternative [...] data has the potential to reveal hidden risks, such as concentrations due to geography or supply chains, and make a significant difference to risk-reward calculations"[12]. But this approach can be applied to any asset class. Let's take an example of a hedge fund manager that operates in the credit space – alternative data, such as "consumer spending" and "payments data," can help with risk mitigation of high-risk borrowers (i.e. borrowers which have a higher likelihood of default than others) by highlighting missed payments and spending patterns.

11 See: https://caia.org/sites/default/files/014-031_monk_jfds.pdf and https://www.risk.net/regulation/6972571/using-alternative-data-to-spot-esg-risks
12 See: https://advantage.factset.com/revolutionising-risk-analysis-with-alternative-data-webcast

Moreover, as we mentioned previously, alternative data can also help hedge fund managers with what is known as operational alpha. In practical terms, this means using unconventional data to improve different business processes[13], such as due diligence or structuring various teams/divisions. For example, novel collation and synthesis of documents can uncover important information about how to improve internal communications, culture and time allocation of human capital, and it can also offer a more detail overview of individual activities of each team's members. Nevertheless, this kind of usage of alternative data is not yet widespread across the industry.

"Alternative data can also help hedge fund managers to manage risk – both investment risk and operational risk."

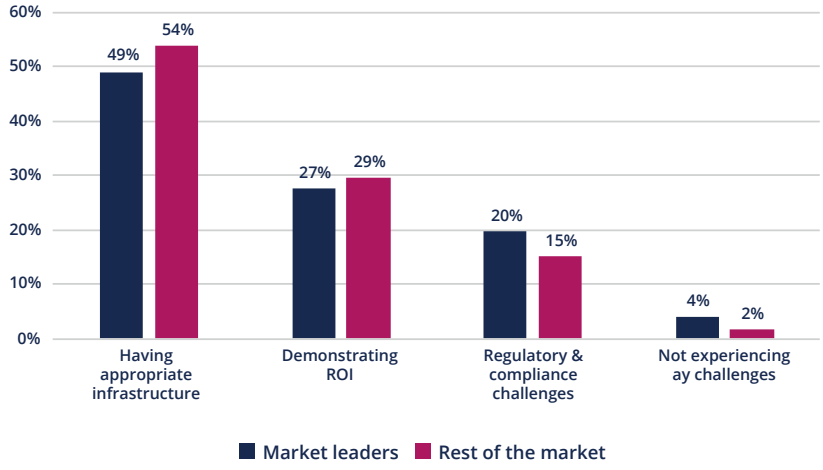
13 See: https://caia.org/sites/default/files/014-031_monk_jfds.pdf

CHAPTER 3

THE CHALLENGES

Working with alternative data in an efficient manner (i.e. in a way to produce either investment or operational alpha) requires having in place the appropriate infrastructure (talent, as well as financial resources), navigating the regulatory environment that governs the collection, usage and distribution of data and demonstrating to investors that alternative data does indeed deliver an attractive return on investment (ROI).

THE CHALLENGES OF USING ALTERNATIVE DATA (USERS ONLY)



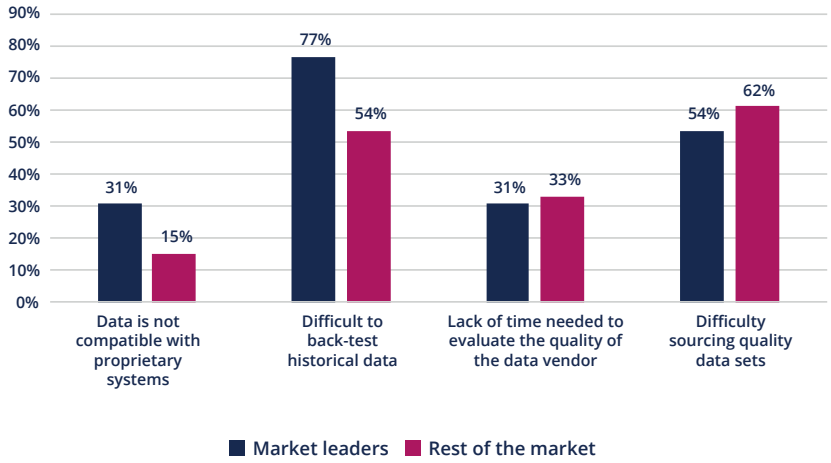
A simple glance at the data illustrates that market leaders seem to have in place a better infrastructure to work with alt data, but they are experiencing more challenges around the nuances of various regulations governing this space than the rest of the market do. Also, unsurprisingly, there are more market leaders, as a percentage of the sample, then late adopters that aren't facing any challenges in this space, suggesting that some hedge fund managers may have figured out the secret sauce of collecting, cleansing and using alt data. Below we look in more detail at the above headwinds.

Building the appropriate infrastructure

Having the right talent and technology – as well as the necessary financial resources to obtain alternative data, backtest it, analyse it and extract useful insights are the pillars of what constitutes an appropriate infrastructure in this context. From our conversations with hedge fund managers, we learned that receiving, storing and consuming alternative data requires a nimble mindset and, in some cases, a different approach than what is required in engaging with traditional data.

Some of the headwinds that our respondents are confronting when trying to build an appropriate infrastructure can be seen in the chart below – notably, market leaders find backtesting alternative data as the biggest challenge. Indeed, the universe of alternative data sets is expanding so quickly that many of them are not going back in time far enough for models to reveal patterns or capture signals. As a CTA manager with about $175 million AUM shared with us, reliability of alternative data and the lack of resources (time, talent and financial capital) are the key barriers that fuel scepticism around the value proposition of alternative data.

CHALLENGES IN BUILDING AN APPROPRIATE INFRASTRUCTURE FOR ALT DATA



Sourcing quality data sets is the second biggest challenge for both user groups; there are already 1,200 datasets in Eagle Alpha's database and the provider estimates that by the end of 2024 there will be about 5,000 alt data sets globally[14]. It is therefore not surprising to see that hedge fund managers find it difficult to know which data sets to look at: being able to focus on what matters while also keeping an open mind for new opportunities is an ongoing balancing exercise that managers have to master.

"Personal experience and proprietary systems play a large role in our evaluation of providers and datasets. While 10 years ago the challenge was to find vendors and data sets to work with, this problem is now flipped on its head. By some accounts there are over 5,000 different datasets offered to the buy-side, so we need to create a process to avoid the bottleneck, evaluate and ingest the best sources of data with respect to return on investment and time. We have evaluated nearly 700 different datasets over the years and use experience as an initial screen before selecting a handful of sets we want to bring into systematic backtesting and evaluation."

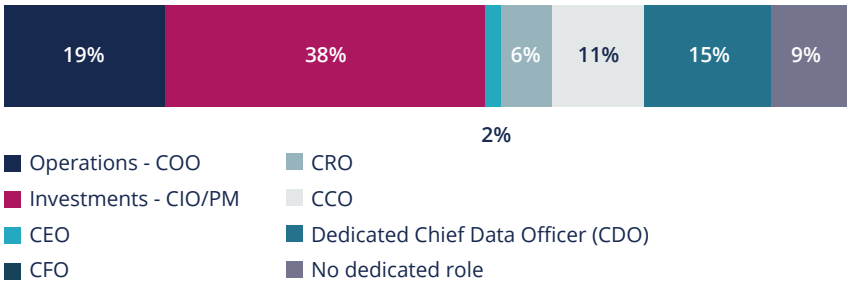
Lombard Odier Investment Managers, 1798 Alternatives

Additionally, even after identifying what data sets to analyse, there are a number of aspects that hedge fund managers need to consider before using the data. These include data governance matters, the breath of data coverage, the robustness of its source, permanency risk (how far into the future can they expect to be using that data set) and how malleable that data is (i.e. whether it can be mapped with fixed references, such as SEDOLs or CUSIPs). Such considerations are particularly important when one considers that alternative data sets are usually discovered as a result of investment research. As a large quant-focused hedge fund manager (with $115bn of AuM) highlighted to us: almost 80% of new alternative data is driven by the research they are doing to test different investment theses, while the remaining 20% typically comes as a result of actively looking for new data sources to analyse and extract novel insights.

To either test an investment hypothesis or gain new information, hedge fund managers will need to hire the right people and to invest in technology such as natural language processing, machine learning (ML) and broader artificial intelligence (AI) platforms that are used to transform alternative data into practical insights.

The chart below illustrates that, for now, the investment department of hedge funds is generally responsible for overseeing the use of alternative data – only 15% of users have a dedicated Chief Data Officer to provide oversight for alt data. However, things are in the process of changing, with more data-minded roles being sought by hedge funds.

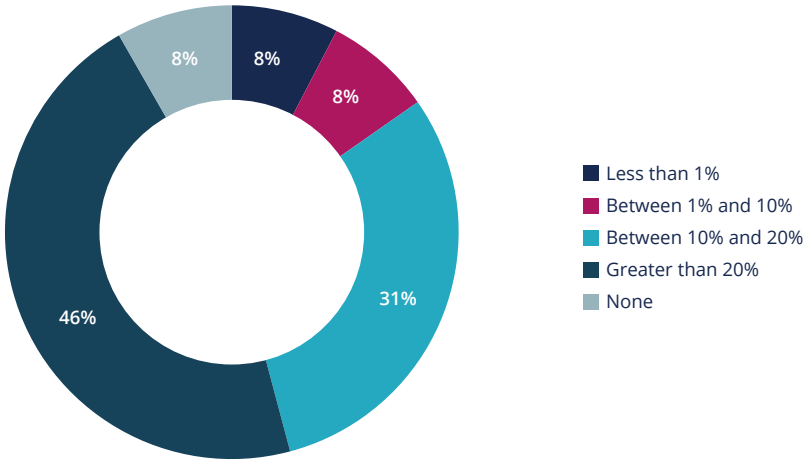
THE ROLE PROVIDING OVERSIGHT TO YOUR FIRM'S USE OF ALTERNATIVE DATA (USERS ONLY)



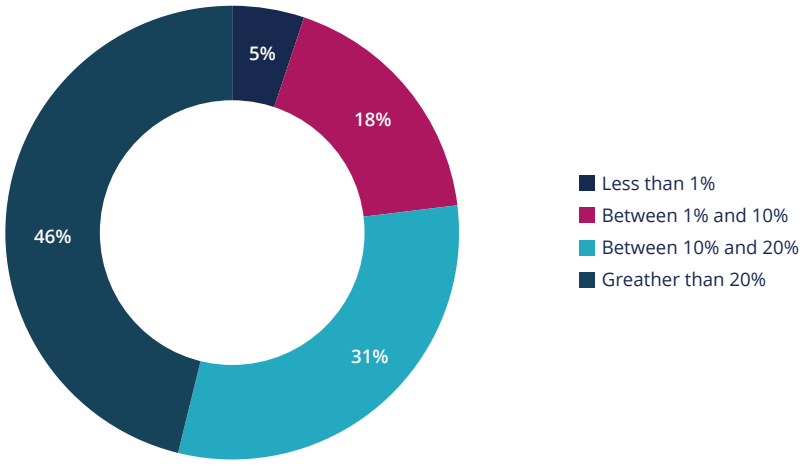
Upon closer examination of the "market leaders", 46% of them are allocating more than a fifth of their investment teams' time and effort towards working with alternative data and 31% of them are allocating between 10% and 20%. These numbers demonstrate their commitment as early adaptors of alt data. Additionally, to continue developing the right talent mix, almost all of them (85% more precisely) are looking to hire data scientists and half of them are also searching for data engineers. Indeed, a recent report from Greenwich Associates states that "data scientists [are] becoming the new quantitative hedge funds"[15].

15 See: Top 10 Market Structure Trends for 2020 - https://www.greenwich.com/market-structure-technology/top-10-market-structure-trends-2020?_cldee=YWJhbGludEBhaW1hLm9y-Zw%3d%3d&recipientid=lead-a38be6641df0e9118117005056ab6b69-1314654de1d943c9ac6e-ba6d96b4a0c3&esid=7ea226c8-9130-ea11-811b-005056ab452b

PERCENTAGE OF INVESTMENT TEAMS WORKING WITH ALT DATA - MARKET LEADERS



PERCENTAGE OF INVESTMENT TEAMS WORKING WITH ALT DATA - REST OF THE MARKET



Similar trends can be seen when looking at "the rest of the market" sample – 46% of the managers falling in this category are dedicating more than 20% of their investment teams' time towards alt data and about a third of them are allocating between 10% and 20%. Data engineer and data scientists also remain, by far, the two most sought-after new hires for the late adopters of alternative data.

Assembling the right teams is fundamentally important in order to extract the benefits and minimise the risks of using alternative data. Indeed, through our conversations with hedge fund managers, we've learned that alternative data requires different teams within the business to work in sync – this means that investment teams, data scientists, risk professionals and compliance and legal teams all need to work together from the point of scouting for new alternative data sources through to the implementation of the insights extracted from them into the investment process. "All groups have to be part of one department working towards common goal", we were told. Therefore, what is really needed is to change the cultural approach to how data is traditionally handled within hedge fund firms: When dealing with alternative data, managers need to understand that the "landscape for generating ideas is changing around them and gaining consistent value from alternative data will require dedication to the project and adjustments to their legacy investment process and team structure", as 1798 Alternatives, which runs a number of hedge fund strategies, including an alternative-data-driven market neutral strategy, told us.

However, the debate regarding what is the right talent mix for working with alternative data will likely continue well into the future, as the alt data sector expands and the hedge fund industry evolves to take advantage of all the opportunities it presents.

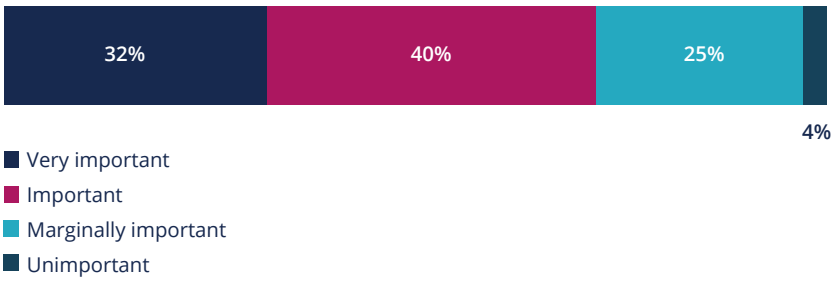
"The landscape for generating ideas is changing around them and gaining consistent value from alternative data will require dedication to the project and adjustments to their legacy investment process and team structure"

1798 Alternatives

Demonstrating return on investment

Another challenge for managers represents the need to demonstrate a return on investment (ROI) on the cost of alternative data. Indeed, 72% of all users (i.e. both "market leaders" and "the rest of the market") say alt data is very important or important to enhancing their overall value proposition. However, justifying the costs associated with alternative data and the advantages hedge fund managers aim to extract from it is key for the wider adoption of alt data sources across the industry.

THE IMPORTANCE OF ALTERNATIVE DATA FOR ENHANCING YOUR VALUE PROPOSITION (USERS ONLY)

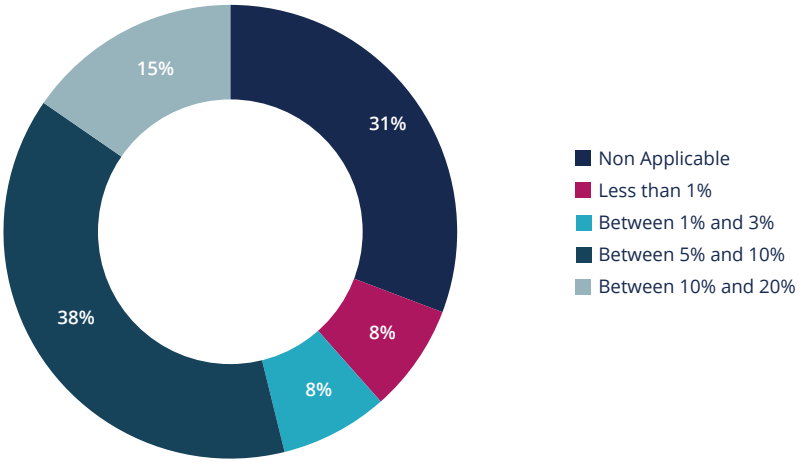


If we look at what our survey data says, the cost of alternative data is relatively small in the context of AUM, suggesting that its cost is not necessarily a barrier to entry: the sample of "users" manages about $382 billion and they are spending around $250 million on alternative data[16], which represents just 0.1% of their aggregated AUM or just 4.4% of the revenues generated through an annual management fee of 1.5%. From this figure, around $103 million is spent on alt data by "market leaders" (41% of total spent on alternative data by the users in our survey), while "the rest of the market" sample is spending $147 million.

16 We reached this figure by assuming an average 1.5% management fee on $382 billion AUM and then calculating the weighted average of expenses (excluding salaries) on alt data.

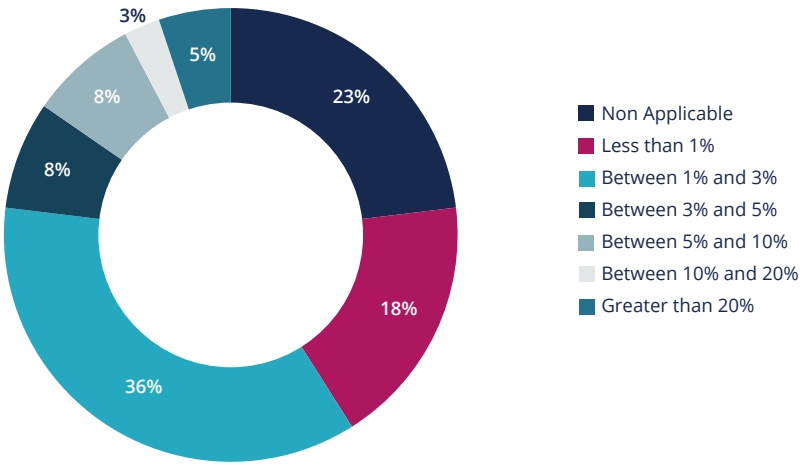
The vast majority of "market leaders" are spending between 5% and 20% of their expenses (excluding salaries) on working with alternative data. Additionally, 83% of these early adaptors have suggested that they plan to increase their spending on alternative data over the next 3 years.

HOW MUCH MARKET LEADERS ARE PAYING TO WORK WITH ALT DATA, EXCLUDING SALARIES, AS % OF TOTAL EXPENSES



Meanwhile, the majority of respondents from the rest of the market sample are spending 3% or less of expenses (excluding salaries) on alt data. We also observe that 5% of late adopters seem to want to catch up with the market leaders and they are spending more than 20% on alt data. At the end of the day, those managers that don't adopt to new ways of generating investment alpha or improving their businesses operationally will not stay competitive.

HOW MUCH THE REST OF THE MARKET IS PAYING TO WORK WITH ALT DATA, EXCLUDING SALARIES, AS % OF TOTAL EXPENSES



However, only looking at expenses from the angle of AUM gives a misleading picture: We need to understand if alt data generates returns that, in dollar terms, are in excess of expenses.

The way hedge fund managers develop ROI measures for assessing whether alternative data is worth the costs depends on whether the manager employs a discretionary or quant strategy. In the case of the former, if the discretionary manager absorbs the costs of alternative data, then they are more likely to make an objective assessment of its value. It is nevertheless difficult to do so, given that discretionary asset management rests on the manager's belief regarding where the best opportunities are and how to go about exploiting them.

That said, quantitative hedge funds are more likely to try to measure ROI from alternative data via an attribution-type of analysis. For example, a large quantitative hedge fund told us that "before onboarding any alternative data, we set a hurdle that requires the strategy using the data to produce alpha net of the costs involving in acquiring, cleaning, structuring and analysing the data". As 1798 Alternatives says, "we target at least a 10x return relative to the all-in cost of any new alternative data sets, substantiated by proprietary software and analytical tools." These players are therefore more inclined to quantify the performance they are getting from alternative data rather than rest on qualitative factors to justify its ROI.

Consequently, it seems that, at this stage of adoption, the question of ROI still remains a mix of qualitative and quantitative equations: In a sense, investors want their managers to be at the forefront of technological change and maintain a step ahead of the competition by investing in new ways of generating alpha. For those managers that run quantitative strategies, measuring the alpha extracted from alt data is easier as they can calculate where each basis point of performance comes from. It is however more difficult to do so for discretionary managers where high conviction plays a greater role in selecting asset class exposure than precise mathematical formulas do.

Indeed, the unique nature of alternative data, as well as the current lack of standardisation around best practices concerning its use, adds extra layers of difficulty in this respect. Therefore, closely related to the absence of robust guidance around alt data is the current state of regulation in this space.

We target at least a 10x return relative to the all-in cost of any new alternative data sets, substantiated by proprietary software and analytical tools."

1798 Alternatives

"We have a data policy in place which describe the data compliance procedure, we measure the 'return on investment' from using alternative data exactly like we measure the efficiency of our predictors, which can be done daily. We follow the allocation process: What is the fraction of the risk of the portfolio that is allocated on signals using a given data set."

Charles-Albert Lehalle, Head of Data Analytics, CFM

Regulatory and compliance challenges

Concerns around obtaining material non-public information (MNPI), data owner consent, privacy issues, intellectual property rights infringement, consumer protection and practices that could provide an unfair advantage are just some of the considerations that hedge fund managers and their legal and compliance teams must account for when using alternative data.

To be clear, it is not the absence of regulations governing the use of data in general that poses a challenge but the fact that there isn't a commonly agreed framework which specifically touches on alternative data is what is causing some caution around adopting it. Indeed, this is a key factor which prevents some of the hedge fund managers with whom we've spoken to adopt alternative data. We are now witnessing an emergence of important legal and regulatory jurisprudence that is likely to start filling in the gaps in understanding of how existing laws apply to the purchase and use of the data. Courts as well as regulators will need to balance the need for ownership rights, consumer and privacy protection with the need to maintain a competitive and innovative digital economy.

From our discussions with alternative data providers, we learned that certain data sets are seen as more problematic from a regulatory perspective than others. For example, some (US-based) alt data vendors use camera images to deliver bespoke analysis of the number of trucks and their loads passing on certain public roads across the USA. To install the cameras and use these images they require permission from local authorities. However, some managers we've spoken with expressed concerns around such bespoke analysis as it may be perceived as an unfair advantage, not to mention their worries about potential data privacy issues.

In a similar vein, providers that derive data sets from tracking mobile information like downloads and user engagement, as well as those that look to capture social media information, can be found be in a difficult spot when it comes to certain regulatory requirements, especially in the wake of the Cambridge Analytica scandal in 2018 or even the more recent case which saw the state of Illinois charge Facebook $550 million for misuse of user data for facial recognition technologies.

"Biometrics is one of the two primary battlegrounds, along with geolocation, that will define our privacy rights for the next generation." attorney Jay Edelson quoted in the Wall Street Journal, January 2020[17]*.*

In particular, hedge fund managers have expressed concerns around exclusive data sets – although having access to alternative data which other market participants can't access may seem as an attractive proposition at first, the potential compliance risks under insider trading and competition law[18] (both domestically and internationally) make it less appealing, not to mention the reputational damage that can potentially ensue from such a situation.

"At CFM, we have a data policy in place which describe the data compliance procedure. Additionally, the CFM compliance team ensures that the data provider has passed the firm's data due diligence process including the answering of the firm's data due diligence questionnaire. As an example, we will take into account whether the provider has the right to collect and distribute the data first. Then one must check if hosting the data in a given jurisdiction can break something (like GDPR if you move the data from the US to Europe). The evolution of these legal aspects has to be monitored even once the data have been bought. In doing so, we would consider a number of things, including if the provider adds a new field a few months later that breaks the anonymity of the data set or if the regulation evolves."

Charles-Albert Lehalle, Head of Data Analytics, CFM

17 https://www.wsj.com/articles/facebook-reaches-550-million-settlement-in-facial-recognition-lawsuit-11580347594?mod=article_inline
18 See for example: https://www.nortonrosefulbright.com/en/knowledge/publications/64c13505/antitrust-risks-and-big-data

The world we live in is becoming more and more digitised and, as such, the amount and types of information that hedge fund managers can use either to research investment ideas or to improve their understanding of current portfolio positions will also expand. Consequently, the topic of data governance and compliance is likely to grow in importance in the years to come, and hedge fund managers that want to be one step ahead should consider developing cyber-security systems which are updated on an ongoing basis: In the fast evolving digital realm, the threat of tomorrow will not be like the threat of yesterday.

Finally, another important obstacle which has not been captured by our survey but was mentioned in conversations with hedge fund managers and which is common to all those that want to embark on a path of using alternative data is ***"dedication of time"***, to use a phrase from ***1798 Alternatives***.

As this hedge fund manager put it:

"A well-built alternative data program requires enough time to succeed. We find people are overly optimistic about how quickly they will see returns or benefits from this initiative. We have been doing this for over seven years and have learned from many mistakes. You can run a data set through a theoretical backtest, but it is a completely different ball game when you put it into production. There are many new developments that have never occurred in the limited historical samples and the investment professionals need to adjust for various biases to the data. This only comes with a learning curve that takes several years."

In order to help hedge funds navigate these challenges, we have put together a "to do" list that managers who want to use alt data ought to complete as much as possible. You can find it at the end of the paper.

CHAPTER 4

WHAT DOES THE FUTURE HOLD FOR ALTERNATIVE DATA?

There is a general sense of optimism regarding the rate of adoption of alternative data within the hedge fund industry. Indeed, 61% of "market leaders" expect alt data to become more widely adopted in the next one to five years, which, as one manager put it, "in financial markets terms means 'now'".

This sentiment is shared by the "rest of the market" as well: 82% of respondents within that sample expect alt data to become more popular within the industry within one to five years. Our discussions with hedge fund managers further confirmed the survey responses – depending on how useful alternative data, in aggregate, proves to the industry, it will probably become more widely adopted faster than most commentators anticipate.

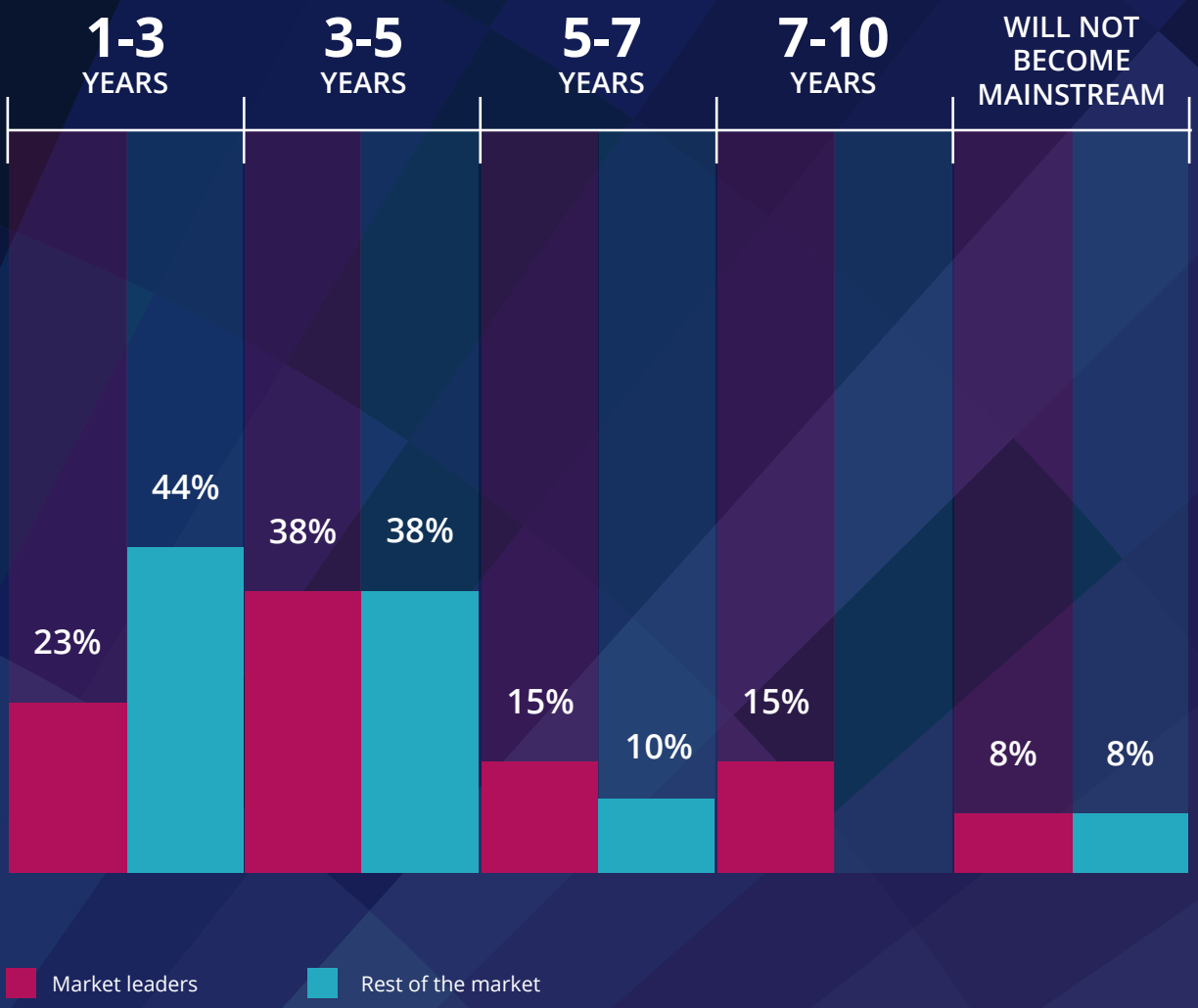
However, in order for this adoption to happen, there are a number of criteria that hedge fund managers believe need to be satisfied. As the chart below illustrates, market leaders and rest of the market have different views in terms of what factors matter most for alternative data to grow in popularity. For the former group, having the required technological capacity to collect, store and assess this type of data is the most important factor – the market leaders seem to understand that even if you have the talent and financial resources necessary to work with alt data, without the software and computational power needed to extract useful insights, the ROI on alt data investment declines.

61% of "market leaders" expect alt data to become more widely adopted in the next one to five years, which, as one manager put it, "in financial markets terms means 'now'".

We believe that alternative data has become mainstream in its use across the hedge fund sector. Just look at the success of conferences like BattleFin, Eagle Alpha or NeuData.

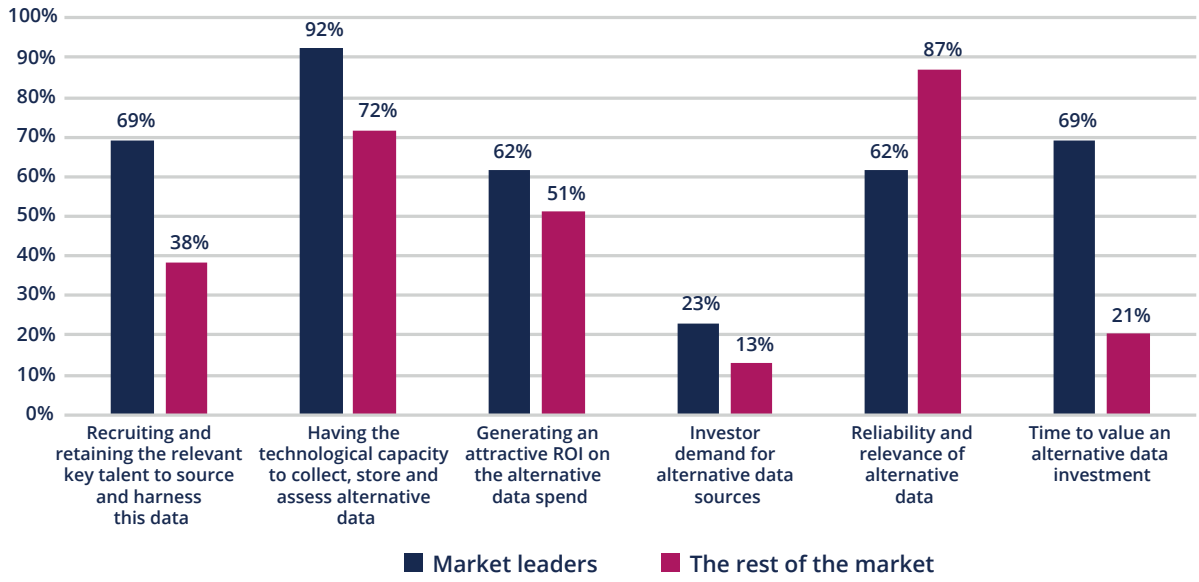
Charles-Albert Lehalle, Head of Data Analytics, CFM

TIMEFRAME FOR WIDESPREAD ADOPTION OF ALTERNATIVE DATA



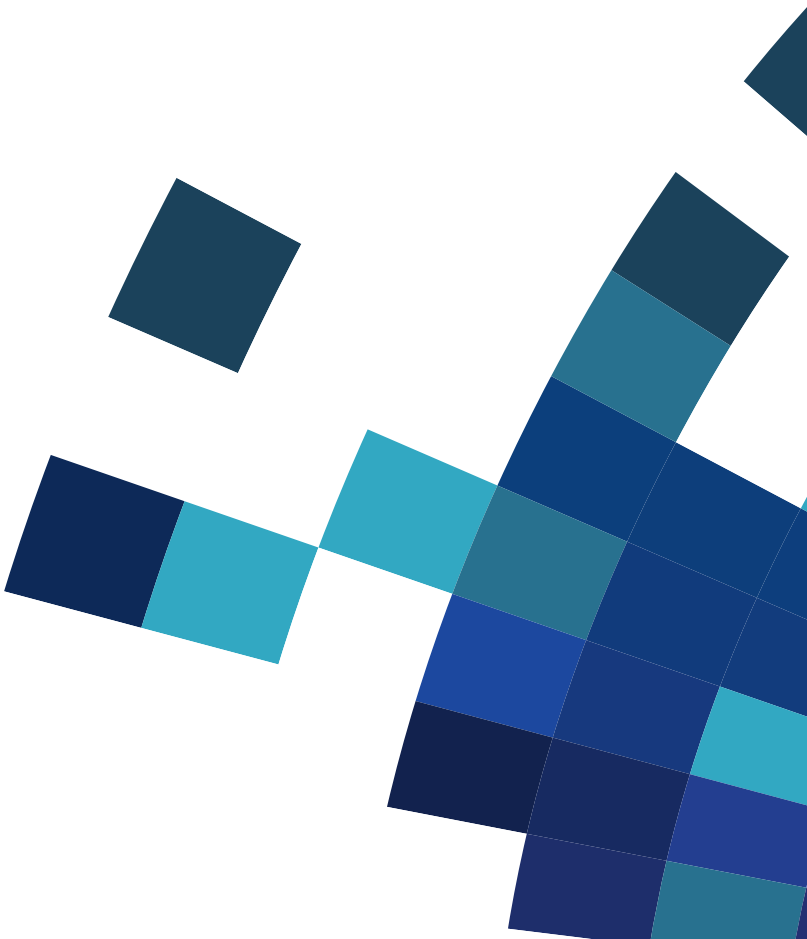
Meanwhile, for the rest of the market, the most relevant factor is the reliability and relevance of such data. This highlights that those coming later on to the party still have to figure out which data sets to focus on. Interestingly, "investor demand for alternative data sources" features quite low on the list of key factors, suggesting that there is a substantial need for educating investors on the opportunities and risks that alt data presents.

KEY FACTORS FOR THE SUCCESSFUL ADOPTION OF ALT DATA ACROSS THE HEDGE FUND INDUSTRY



Adding further colour to the above data points, our discussions with managers revealed that for alternative data to become more mainstream both vendors and users need to be pragmatic and flexible and, through clear and open communication, collaborate on building the right data sets that deliver an attractive ROI on its costs – translating alternative data into actionable insights requires unique skillsets, which can represent a significant barrier to entry. However, everything has a price. Therefore, the costs associated with using alternative data need to be justified for its adoption to become more widespread.

Despite much of the exciting progress within the field of Big Data (and consequently, of alternative data) which we have witnessed over the past decade, we are only in the very early days of such developments. The next ten years will likely see even more growth of alt data, with some managers expecting the next five years to bring four-to-five-times growth in this space. Overall, hedge fund managers are optimistic about the growth prospects of alt data within the industry, with many of them planning to hire more people, dedicate more of their investment teams' time on it and also spend more on this type of data. Nevertheless, the process of adopting alternative data won't be a smooth road: A "trial and error" approach is likely to be the common path that most hedge funds will embark upon.



PRACTICAL STEPS FOR HEDGE FUND MANAGERS LOOKING TO USE ALTERNATIVE DATA

Some of the experienced alternative data users have developed strong policies and processes around onboarding and using alternative data in order to minimise the risks involved.

A large quant hedge fund explained to us their process. It starts with an initial interview of the alt data provider, focusing on items such as the source of the data, how the vendor deals with privacy and MNPI issues and whether appropriate legal protections for clients are in place. Following this stage, a further interview with the data provider takes place where a compliance professional is in attendance as well. Then, a request to complete the due diligence questionnaire that AIMA has designed is sent to the vendor. For trial models, the Standards Board for Alternative Investments' "Big Data: Standardised Trial Data License Agreement" is usually used in order to address key issues such as personal data protection, prevention of insider trading and the "right to use data". The final step is to a get a production licence, preparing the data to be incorporating into the investment process.

Although there are general steps that all hedge fund managers need to take when using alternative data (for example, ensuring compliance with the relevant laws and regulations), each manager is ultimately different in how they decide to use alternative data. Below we provide some generalised practical steps for managers looking to use this type of data. These steps have been compiled based on conversations with hedge fund managers that are already working with alternative data or that are looking to do so.

Alternative Data: Exploring beyond the realm of traditional finance and economics

- Ensure that you have the right infrastructure in place before deciding to start onboarding alternative data. This means having the necessary technology and talent that can work with this kind of data. Some considerations include:
  - Hiring data scientists and engineers
  - Building in-house digital platforms (or using off-the-shelf options) that enable interaction with alternative data, including software tools and statistical concepts provided by machine learning – this will help with comparing the efficiency of data sets based on the same economic rationale but provided by different vendors
  - Identifying third parties to which parts of the process (e.g. data cleansing) can be outsourced
  - Drafting policies governing the use of alternative data – quantitative and qualitative due diligence are crucial
- Once the appropriate infrastructure exists, ensure that all the departments which are going to be involved in the process of onboarding, cleaning, assessing and making decisions based on alternative data are in sync. Things to bear in mind include:
  - Open and clear communication between departments is key
  - Establishing a workflow that ensures smooth onboarding and subsequent use of alternative data

- Conduct thorough due diligence on the alternative data vendor using AIMA's due diligence questionnaire available **here**. This will ensure that the alt data is sourced from a reputable provider. Some of the key things to consider at this stage are:
  - How does the vendor procure the data?
  - Are there any concerns around the governance of such data?
  - How reliable are the sources from which the alternative data in question comes from?
  - What are the risks associated with data discontinuity in the future?
- Embark on a trial period following the Standards Board for Alternative Investments' "Big Data: Standardised Trial Data License Agreement".
- Conduct a cost-benefit analysis to ensure the ROI on alternative data is justified.
- Catalogue and tag the data in order to increase the chance of discovering new insights and in order to reduce duplication in sourcing and evaluating it.
- Conduct periodic reviews of assessing how impactful alternative data has been in generating an investment edge or operational alpha: Focus on nowcasting first, then on forecasting.

CONCLUSION

This paper provides an overview of how the hedge fund industry is using alternative data to generate returns and mitigate different forms of risks. As a concept, alternative data is not new – indeed, it is quite ancient.

Investors have always been innovative in seeking out new ways to gain an informational edge. History is full of examples demonstrating this, from ancient Babylon merchants using measurements of Euphrates' depth and flow to inform their decisions in trading various commodities through to Venetian traders' use of telescopes to inspect the flags of incoming ships in order to derive what type of cargo was being carried.

Data is a big disrupter now. 90% of the world's data available today was only produced over the last two years – a truly staggering statistic. The pace is only likely to accelerate more, so the opportunity set for investors globally is compelling.

Alternative data is proving to be useful in moments like we are currently experiencing with the recent global pandemic. The virus has led to a surge in demand for alternative data from investors, governments and central banks to best interpret how the pandemic is impacting industries around the world. Purchase receipt data from credit card transactions, geolocation data derived from satellite images have all been used to help investors get a real time insight into how economies are reacting to the disruption. The immediacy of these data sets in comparison to the information lag from working with traditional data is particularly helpful in moments like this, when markets cease to function normally.

Hedge funds have always embraced change, constantly innovating in order to deliver superior returns to investors. Many of the industry's longest established names have been at the forefront of rigorous data driven investing.

The confluence of more data becoming available each day and the rapid leaps being made in technology can only improve their decisions making - whether it is through using these new forms of trade to inform a human decision or for them to serve as inputs for an algorithm which can then make a trade.

With a growing number of alternative data providers, hedge funds have access to a wide range of data sources. Throughout the current disruption we have seen examples of hedge fund managers successfully employing alternative data in navigating market conditions.

One Asia-based manager used alternative data to understand how markets may react and, importantly, how to position the portfolio. In particular, it combined information on past outbreaks with web crawled data which contained internet searches and insights about road congestion, flight schedules and test-kit availability. The manager subsequently delivered outstanding returns to investors as markets turned more volatile and declined considerably during the first months of the pandemic.

Throughout this research, there is broad optimism regarding hedge funds continued use of alternative data in the years ahead, with many of our respondents believing that its use will become more mainstream soon. The challenges being presented to hedge funds and to industries more broadly are how best to gather this type of data and more importantly, how to derive something meaningful from it. Hedge funds with their long tradition of pioneering change and bringing new forms of innovation to investment management are ready to take up this challenge.

About AIMA

The Alternative Investment Management Association (AIMA) is the global representative of the alternative investment industry, with more than 2,000 corporate members in over 60 countries. AIMA's fund manager members collectively manage more than $2 trillion in hedge fund or private credit assets. AIMA draws upon the expertise and diversity of its membership to provide leadership in industry initiatives such as advocacy, policy and regulatory engagement, educational programmes and sound practice guides. AIMA works to raise media and public awareness of the value of the alternative investment industry. AIMA set up the Alternative Credit Council (ACC) to help firms focused in the private credit and direct lending space. The ACC currently represents over 170 members that manage $400 billion of private credit assets globally. AIMA is committed to developing skills and education standards and is a co-founder of the Chartered Alternative Investment Analyst designation (CAIA) – the first and only specialised educational standard for alternative investment specialists. AIMA is governed by its Council (Board of Directors).

For further information, please visit AIMA's website, www.aima.org

About SS&C

SS&C Technologies Holdings (NASDAQ: SSNC) is the world's largest hedge fund and private equity administrator, as well as the largest mutual fund transfer agency. SS&C's unique business model combines end-to-end expertise across financial services operations with software and solutions to service even the most demanding customers in the financial services and healthcare industries. SS&C owns and operates the full technology stack across securities accounting, front-to-back-office operations, performance and risk analytics, regulatory reporting, and healthcare information processes.

SS&C's trusted and proven technology delivers an unparalleled level of scalable capabilities for the most complex portfolios, the most sophisticated strategies, and the highest volumes of transactions. Through a series of carefully selected acquisitions and organic growth, the breadth and depth of SS&C's expertise in financial services and healthcare technology are unmatched.

For further information, please visit SS&C's website, www.ssctech.com

AIMA
Alternative Investment
Management Association

AIMA.ORG