2022 Alternative Data Report: A Year in Review

December 2022

www.eaglealpha.com
# Table Of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Letter From Niall Hurley, CEO</td>
<td>3</td>
</tr>
<tr>
<td>Section 1: A View from the Front: A Data Sourcing Perspective</td>
<td>6</td>
</tr>
<tr>
<td>Section 2: A Synopsis of Mergers &amp; Acquisitions in the Alternative Data Ecosystem</td>
<td>12</td>
</tr>
<tr>
<td>Section 3: A Year in Thought Leadership: Trending Content from our Premium Data Strategy Portal</td>
<td>18</td>
</tr>
<tr>
<td>Section 4: Leading Case Studies Of 2022</td>
<td>25</td>
</tr>
<tr>
<td>• Inflation and Financial Markets</td>
<td></td>
</tr>
<tr>
<td>• The Volatility of the Crypto Markets</td>
<td></td>
</tr>
<tr>
<td>• Analyzing China’s Slowdown in Trade and its Potential Beneficiaries</td>
<td></td>
</tr>
<tr>
<td>• The Latest on Layoffs: Job Listings Data Provides Early Warning Signs</td>
<td></td>
</tr>
<tr>
<td>• The Rise of Alt Data Adoption in Private Markets</td>
<td></td>
</tr>
<tr>
<td>• Natural Language Capabilities</td>
<td></td>
</tr>
<tr>
<td>Section 5: The Year in Market-Driven Product Development</td>
<td>35</td>
</tr>
<tr>
<td>Section 6: 5 Key Predictions for 2023</td>
<td>41</td>
</tr>
</tbody>
</table>
Dear All,

Welcome to all our readers to the Eagle Alpha Annual Report for 2022. It has been another dynamic year in alternative data, and we look forward to updating you on some of the standout themes and developments both on the Eagle Alpha platform and within the broader marketplace.

Speaking first to marketplace growth, the positive dynamics relating to supply growth of new data sources continue with structural trends around data creation and productization that have not changed. We continue to add approximately 20 data products to the platform every month despite the hurdles to getting onto the platform becoming more stringent with the completion of an FISD standard due diligence questionnaire (DDQ). Within this report, we will discuss some of the more granular category and geographic trends we have seen.

Corporate data monetization is still an important topic. There is a view that some of the most interesting data products are still to come with internal data assets at corporates being made available for external data consumers. Such exercises are complex and multiyear in nature and involve significant alignment of stakeholders within a corporate environment. The positive news is that these conversations are prevalent, and corporates are engaging with Eagle Alpha to understand the value and relevance of their datasets while ensuring they address compliance and privacy in their productization journey.

While the structural trends around data adoption remain robust within funds, private equity, consultants, and corporates, this year brought significant disruption to many investment strategies. The drawdown across asset classes in capital markets due to geopolitics and interest rates were as severe at one point in 2022 as in the 1930s. Inflation has eroded the value of cash and everything from equities, bonds, commodities and crypto currencies have received draw downs.

This meant that the divergence of strategy performance was particularly pronounced versus history. Commodity Trading Advisors (CTAs), macro, multi-strategy and certain quantitative strategies have thrived in these market conditions and their ongoing appetite for data has not abated. For other firms that are long-only or directional in nature, it has been a more conservative year with performance and management fees under pressure. This was challenging in the first half of 2022; however, the dynamic improved in the second half.
Private equity remains an important vertical for Eagle Alpha and the data community. We held our first private markets dedicated conference in Boston in June 2022. We had the opportunity to hear from specialists at Bain & Co, PwC, Bain Capital and many others on their perspectives regarding data. This is a vastly different part of the marketplace with data needing to integrate into due diligence, deal origination, and portfolio company exercises in a quite different manner to public markets. There are solutions to this need and Eagle Alpha is working closely with data consumers and vendors to help drive innovation to support adoption here. As I write, the financial press are looking at the recent turbulence in the crypto markets and asking if sufficient due diligence was applied to capital allocation decisions. It is difficult to see how due diligence will not evolve further post this episode. Alternative data has a role to play.

One of the most important developments at Eagle Alpha this year has been the evolution of our platform in relation to data product insights and compliance. The platform features became more relevant due to the SEC Risk Alert in April which provided guidelines to compliance teams at investment advisors. While investment firms are strong at compliance at ‘onboarding’, the SEC Exams Division said that compliance teams needed to improve their ‘re-profiling’ and monitoring of risks relating to dataset compliance during the data lifecycle.

The SEC Exams Division found during the subscription that there were not sufficient risks or alerts if there were issues or changes with a data vendor. We are solving for both callouts with time-stamped answers provided by the data vendors and notifications for changes. Next, we want to capture communications.

It should be no surprise that the compliance team is now an important client for us in both funds and private equity that are governed by the Investment Advisers Act.

Importantly, the pandemic environment and severe restrictions on in-person engagement lifted this year. Conferences remain an important gateway into the Eagle Alpha ecosystem for buyers and vendors and the team was delighted to see clients and new contacts in New York, Boston, and London in person. Our events go from strength to strength and are more integrated into our platform than ever.

All this means that Eagle Alpha is built for scale and all the platform functionality is becoming more self-service in nature for both the buyer and vendor and different personas within both groups. As we progress towards 2,000 or even 3,000 data products with more functionality moving beyond profiling, to deeper compliance, connections, and data delivery, the importance of being built for scale and operational efficiency is key. Our long-standing clients are utilizing more functionality and driving our platform development based on their workflow needs.

Finally, I wanted to say a significant thank you to the entire team at Eagle Alpha. In the last 2 years, we have achieved a lot. We have focused our client offering on 2 solutions, ramped platform functionality with 8-week development cycles, and brought new talent into the business, with Eagle Alpha team members now present in Ireland, the United States, the United Kingdom, and Spain. Importantly, we have managed to navigate a pandemic and bring the business forward into new verticals.

Thank you to all our clients on both the buyer and vendor side of the platform – we have worked with you for many years and your energy, feedback and commitment to our partnerships are fully appreciated. We look forward to seeing many of you on January 19th in New York.

Regards,
Niall Hurley, CEO
Beyond our platform and marketplace developments, I wanted to reflect on last year’s predictions and the year ahead. In December 2021, we predicted the following events to take place in 2022:

1 The ESG trend is not going away – this has certainly played out with ESG remaining key but the regulator is becoming very focused on how funds are presenting themselves at ESG versus their underlying research and portfolio practices.

2 European transaction data would receive more attention – this has certainly been the case with Consumer Edge, a category leader, acquiring Qentris, and Fable Data looking to expand its product offering with increased country coverage.

3 KPIs for working with data are going to improve further – we have witnessed this with a growing number of our clients moving to master trial documentation, standardized paperwork, and improving their overall efficiency. We believe compliance-related KPIs will be next.

4 Data-driven funds are going to win the AUM race – the dispersion we saw this year between AUM and the performance of data-driven quants and multi-strategy firms and the race for capital, data, and talent remained significant.

5 How buyers and vendors engage is going to get smarter – we now have compliance teams engaging directly with vendors via Eagle Alpha. Vendors can prioritize buyers more intelligently via our platform. Look for direct communication capture for compliance teams within our next development cycle.

6 The app engagement category would be active – while the category remained active, we probably did not see the level of churn or price volatility that we expected one year ago.

7 The compliance team is a client alongside the data team – this was very prevalent, and we have seen the compliance team lead subscriptions to Eagle Alpha for the first time.

8 Industry consolidation will continue. For more of our insights, see our M&A in Alternative Data starting on page 12.

Our predictions for 2023 will be found on page 46.

How buyers and vendors engage is going to get smarter – we now have compliance teams engaging directly with vendors via Eagle Alpha.
Section 1

A View From The Front – A Data Sourcing Perspective
Celebrating our 10th year as a company in September 2022, we continue to see structural growth in the industry due to ongoing data supply. In 2022 we continued to add approximately 20 data products per month to our platform. Under our new dataset taxonomy you can see the yearly expansion by data category below.

Importantly, dataset growth now also means deeper due diligence on the platform. For every data product added we now add a DDQ. In 2022, that meant 250 data products and 250 DDQ were added to the platform.

We are transitioning and evolving from cataloguing to data connections and data deliveries. This also means that we are increasingly an integrated workflow tool for compliance, researchers and data operations.

Deeper profiling and data connections is a new development in 2022 and is being led by client demand. We are just beginning this journey but in the last 4 months of 2022 we connected 40 data products, with the London conference acting as a catalyst. In time we believe a data connection request from clients will become as natural as a DDQ request.

Note: A dataset can cover more than one category. 1,750 datasets related to 2,100 main categories.
On a regional basis data growth also continued. This is visible from the chart on the left. Most of the growth came from vendors classifying themselves as global or from developed markets. Although not shown here, buyer use of data from China stalled in 2021 and much of 2022 but we have seen some renewed interest in China over the closing months of 2022 from larger funds with global mandates.

The top five growth areas in our platform in 2022 is shown in the plot on the left. Category growth is somewhat distorted this year due to our reclassification of data categories. However, Crypto datasets have seen strong growth as data interest was nascent in 2021 and had very strong interest in 2022. The Financial Data category grew largely on trade signalling or trading model datasets. Interest in ESG has waned this year but dataset growth remained strong in the year.
Throughout 2022, it became clear that certain aspects of the industry are changing. This is leading to new players in the industry and also growth in sources of information. We have seen rapid vendor growth throughout the year and an increased spread of data buyers. A stand-out change on the buyer side is the use of alternative data by private markets and consultant firms.

To address asset manager demand and new buyer interests we added new features to our product offering in 2022. Several new features help buyers source relevant data for their given use case.

Since the launch of the industry taxonomy of 26 data categories in 2016, we have taken note that the categories have changed and evolved to be more comprehensive and reflective of the demand in the market making the case that umbrella categories with more granular subcategories are necessary.

Responding to this, one significant new feature is a reorganization of our dataset taxonomy. While it has just been recently launched, it has been many months in development. The old taxonomy had become stale and difficult to work with.

We now have a new taxonomy of 16 primary categories and 56 subcategories. This transformation was quite complex and took some time for us to implement and validate, but we are certain that our clients will find the new taxonomy useful in the coming years and this new categorization will be adopted throughout the industry. More details on these features can be seen below in the tech product update section.

**Figure 3: Eagle Alpha’s New and Improved Alternative Data Taxonomy**
Client-led requests with questions coming primarily from asset managers and private equity prompted data sourcing to identify the most interesting high-level trends and themes throughout the year. This section highlights these trends.

**Employment**

At the beginning of the year, we had an unusual spike in interest in the employment data category. I use the term unusual here as the employment data category is well-established and deployed across the investor universe. Interest came for both human capital data and job posting data categories (our new sub-categories!). Perhaps, with the great resignation asset managers and PE firms were trying to get a handle on attrition rather than just employment trends. Looking deeper into hiring and attrition at the O*NET code or job role also being an important signal.

Also interesting in this category, we added several new employment data vendors to the platform. The category is quite crowded, and it will be seen how this competition plays out in the coming years. Recent legal rulings may also have an impact on the data vendors in 2023.

**Inflation**

Perhaps not too surprisingly we had a lot of requests for supply chain and inflation data over the course of the year. Inferess, Factset, and China Scope all received a lot of interest from buyers for the supply chain question. For inflation, the answer, such as it is, was complicated and involved a myriad of data options across real estate, travel, and durable and non-durable goods, for example.

This question is more complicated when you look at core CPI or the PCE deflator in the US and HCIP in Europe. While most funds gravitate towards building themselves versus using someone else’s output, the alternative data macro vendors received some interest from inflation and macro signaling point of view.

A side point on this topic is that we have seen a trend for Macro/CTA funds to start looking more at alt data and some larger global funds and multi-strategy funds appear to be expanding into this part of the market.

**Consumer Transaction**

One big theme to hit the alt data ecosystems occurred across the consumer transaction space. From July onwards we were fielding conversations with clients on panel disruption and panel changes across the US consumer transaction providers. Initially, the disruption came from a “channel partner” (Yodlee) to several data vendors such as M Science, the assets Facteus acquired from 1010, Second Measure, and Earnest Research.

The panel disruption was expected originally to be temporary but as summer dragged into Autumn this panel issue was not resolved and currently appears like it will not be resolved. In response Second Measure moved to launch a new data offering with a new data panel, M Science began to emphasize SunDial more and Facteus launched a new offering called Facteus Pulse to replace 1010 Data and any other impacted data vendor.
In October Earnest announced a new product called Vela that is using the same data panel provider that Consumer Edge is using. Through the summer the chatter was that Consumer Edge had lost the exclusive with its data provider and this proved to be the case when Vela was launched. The consumer transaction space remains in flux as all these vendors launch new products and funds are now back-testing these offerings. Consumer Edge remains in a strong position, particularly with the acquisition of Qentnis, but competitively there are more options now for buyers, and pricing trends will be interesting to observe next year and beyond.

Natural Language Processing

We have seen a renewed interest in unstructured data and NLP for seven or eight months. We are not sure what is driving this but the interest in coming from the three areas of company-reported data, transcripts, and general news. Sometimes a collection of the three sometimes just a given topic. The interest is across geographies and unstructured data in local, non-translated, text is important. While the interest has been for unstructured text there has also been an interest in vendors who are supplying the additional layer of NLP, sentiment, and scores at a topic and ticker level.

Regional Considerations

Regionally a good portion of our clients is US asset managers and PE funds. However, we continue to receive requests for information and interesting datasets across the ASEAN and APAC regions. Investors are looking into developed and developing markets in the APAC region but in general developing markets had more of a focus.

Requests on China datasets remain somewhat subdued given regulatory issues. Or sense is that funds may be setting up or expanding operations in China to circumvent the data export restrictions. This may lead to funds with global mandates and infrastructure to reengage on the China data side. Vendors remain hesitant to deliver data to foreign entities.

The consumer transaction space remains in flux as all these vendors launch new products and funds are now back-testing these offerings.

Environmental, Social, and Governance (ESG)

One last theme for the year is a theme from 2021 that has largely expired. Our requests for ESG-related data have been light in 2022. This follows a year in 2021 when it was very prevalent. At the end of last year, we talked to many ESG vendors and they all said they had a busy year in terms of client conversations, albeit with some level of tire kicking. As our line of sight into funds moving beyond talking and testing to buying certain datasets is limited, we do not know if 2021 conversations turned into 2022 purchasing. In any case, our level of engagement this year on ESG has been low.

We look forward to resuming our data conversations with clients and prospective clients next year and hope the new features on our platform help all concerned to find and prioritize their data needs.
Contributors:

Mikheil Shengelia
Research Analyst, Data Strategy

Dallán Ryan
Content & Advisory Lead, Data Strategy

Section 2

A Synopsis of Mergers & Acquisitions in the Alternative Data Ecosystem
In this section, we highlight several mergers in the alternative data industry over the years and present exclusive commentary from industry leaders.

New industries are typically fragmented with numerous small players competing for market share. Industries then consolidate as they become more mature. According to A.T. Kearney’s study, all industries go through a consolidation life cycle which consists of five stages. Every company in every industry goes through these stages or disappears. Figure 1 below shows that the consolidation curve might take different forms, but industries will nonetheless go through the Opening, Plateau, Concentration, Dominance, and Reopening stages.

Sam Livingstone, Head of Data Science at Jupiter Asset Management, commented on the market evolution at Eagle Alpha’s 2022 London Conference: “Data vendors go through a similar hype cycle. They feel that they have data no one else has, then realize that there are others offering similar products. Then they make it expensive, then make it more realistic. This is a trend I see continuing.”

Increased competition, changing regulations, new technologies, and increasing customer sophistication are putting pressure on today’s technology leaders and alternative data vendors. Companies would mostly pursue mergers and acquisitions to gain scale advantages and accelerate growth. A size advantage often lowers costs leading to competitive advantage and competitive price setting.

Ed Lavery, VP of Investor Intelligence at Similarweb, predicts consolidation in each alternative data category: “I do think the larger data buyers will continue to aggregate more datasets to increase their coverage and accuracy of their existing offering.

These same data vendors will continue to widen the moat of their niche, and mass market buyers will therefore have a much more limited pool of vendors to purchase from, each with their own specialty and focus.”

Nasdaq’s Acquisition of Quandl

The exchange group’s decision to acquire Quandl in December 2018 meant that Nasdaq was revamping its overall company strategy and expanding beyond traditional services of listings and into data and technology. Nasdaq’s Analytics Hub included 25 datasets at the time, and it was integrated with Quandl’s platform consisting of 350 alternative datasets. Prior to that, Nasdaq also acquired buyside behavioral analytics specialist Sybenetix, investment analytics provider eVestment Alliance, and trading technology provider Cinnober.

Bjørn Sibbern, executive vice president for Nasdaq’s Global Information Services, explained the rationale behind the acquisition of Quandl: “We want to grow our data and index business, so adding other types of data—in this case, alternative data—is the natural thing for us to do. Firms need smarter data for investment decision-making, and we believe alternative data will become must-have data for the buy and sell sides.”

Nasdaq fully integrated Quandl and later introduced its Data Link product, a centralized platform that provides access to more than 250 datasets. Other notable acquisitions in recent years include Puro.earth, a leading marketplace for carbon removal instruments, and Metrio, an ESG analytics and reporting services vendor.
ExtractAlpha’s Acquisition of Estimize

In May 2021, ExtractAlpha announced the acquisition of Estimize, an alternative data vendor providing crowdsourced financial estimates. The company focused on crowdsourcing EPS and revenue estimates from over 65,000 active contributors: independent, buyside, and sell-side analysts, private investors, and students. The idea is to provide a more accurate and representative view of expectations.

Vinesh Jha, CEO of ExtractAlpha, said: “We will be leveraging the strengths of each company – Estimize’s strong community and unique dataset, and ExtractAlpha’s research and sales expertise – to reach new customers and create exciting new product offerings.”

The Merger of Burning Glass Technologies with Emsi

In June 2021, Burning Glass Technologies announced a merger with Emsi to create a combined entity called Lightcast. The largest investor of Burning Glass, KKR, made an additional investment whilst Emsi’s owner Strada Education Network made an exit. At the time, Burning Glass was already one of the world’s largest providers of job openings data with more than 40,000 daily sources in over 30 countries.

And Emsi’s dataset comprised 170 million employee profiles in the US with detailed histories capturing where employees came from and what companies they are going to, including their education and skill set.

Matt Sigelman. CEO of Burning Glass, said: “Our complementary capabilities in data analytics and product development will give us the scale and resources to deliver even greater value to our customers through new, enhanced products. We are looking forward to working together to continue to drive transformative change in the labor market.”

“The world of work continues to undergo seismic changes as a result of automation and the COVID-19 pandemic and combating the resulting labor market challenges has become an economic and social necessity. We believe the combination of these two industry leaders will be uniquely positioned to address these unprecedented challenges and we look forward to supporting them as they do,” said Kyle Matter, Managing Director of KKR’s Global Impact team in North America.

NielsenIQ’s Acquisitions

Nielsen Holdings spun off its consumer data unit in 2020 which was acquired by private equity firm Advent International. The newly formed company is now known as NielsenIQ and it made two acquisitions in September 2021 – Data Impact and Rakuten Intelligence – to accelerate the growth of its e-commerce and omnichannel measurement solutions.

Clément Colin, Head of International E-commerce Measurement, said: “We are proud to welcome these two champions into the NielsenIQ family. We will be able to provide a single number for online sales measurement through multi-sourced data—the truth the market needs—and we will also deliver e-commerce execution recommendations with unrivaled actionability.”

In order to further expand its coverage and enter new markets, NielsenIQ acquired Foxintelligence, an e-commerce measurement and consumer data analytics company based in Europe. “NielsenIQ’s ambition is to approach e-commerce holistically. Foxintelligence is instrumental in this strategy, expanding the footprint of our unmatched retail measurement solution. They uncover the online blind spots like marketplaces, category specialists, direct-to-consumer, quick-commerce, and last milers, among others, providing full-market coverage,” said Clément Colin.
Similarweb’s Acquisition of Embee

In November 2021, Similarweb acquired Embee, a San Francisco-based mobile insights provider, to extend its user datasets. Embee’s patented technology was designed to provide end-to-end research panel recruitment, measurement, rewards, and retention within a single application.

“Embee’s behavioral data and digital measurement approach — backed by a large-scale metered panel of highly engaged, opt-in users — enables Similarweb to strengthen its data edge in digital measurement and improve its mobile app intelligence offering. The granular insight into mobile app usage that Embee enables is extremely valuable across a wide variety of industries and use cases. The unique Embee panel positions us to continuously introduce new market research capabilities with differentiated mobile insights” – said Ed Lavery, VP of Investor Intelligence at Similarweb.

In an interview with Eagle Alpha, Ed also made predictions about data buyers going forward: “In reality, as alternative data is more widely adopted, investors won’t need so many data sources. The biggest consumers of alternative data today (c.50 funds) will continue to look for variety, but the mass market will look for simplicity. If a handful of data sources can tell 80% of the story, then that will be sufficient for most buyers. Going forward, more and more buyers are going to enter the market, and there is going to be an increasingly large number of investors with limited or no data sophistication. There is going to be an increasing amount of expectation from investors for more of the work to be done by the data vendors – i.e cleaning, finding the signal, integration, etc.”

Comscore’s Acquisition of Shareablee

In December 2021, Comscore announced its acquisition of social media data provider Shareablee and built a taxonomy of over 20 million publishers, advertisers, and content creators. Shareablee captures the social media actions of more than 2 billion consumers and parses them across 700+ metrics for brands, publishers, TV shows, influencers, and celebrities going back to January 2013 on platforms including Facebook, Twitter, Instagram, Pinterest, and more.

“Comscore is committed to the future and helping customers have a comprehensive view of the consumer across platforms while understanding the unique engagement on premium video. The future of measurement would not be complete without including the impressions delivered by the social media platforms,” said Bill Livek, CEO and Executive Vice Chair at Comscore.

S&P Global’s Acquisition of IHS Markit

In February 2022, S&P Global announced its acquisition of IHS Markit to provide combined cross-sector, fundamental, and alternative data with workflows and solutions across its expanded customer base. S&P Global’s marketplace now includes a variety of data sources including bill of lading data, well logs, ship movements, and supply chain trends.

On this acquisition, Douglas L. Peterson, President and Chief Executive Officer of S&P Global said: “Our combined strengths in credit and risk management, indices across multiple asset classes, private markets, ESG, and energy transition data and analytics, will accelerate the growth of our business and broaden the scope of services we can provide to the markets.”
This deal needed to pass the EU Merger Regulation and be approved by the European Commission. However, the approval was conditional on the divestment of businesses in commodity price assessments and financial data. Other notable acquisitions by S&P in recent years include climate risk analytics provider The Climate Service, emerging technology data and intelligence provider 451 Research, and AI solutions provider Kensho.

**FactSet’s Acquisition of CUSIP Global Services**

In March 2022, FactSet completed the acquisition of CUSIP Global Services from S&P Global. The deal expanded FactSet’s role in the global capital markets as it now partners with American Bankers Association which launched the CUSIP system back in 1968. Scott Preiss, Managing Director and global head of CUSIP Global Services, said: “The entire CGS staff looks forward to joining the impressive team at FactSet and to ensuring that the CUSIP system continues to deliver best-in-class securities identification and data management solutions.”

Other notable FactSet acquisitions include Revere Data, an industry taxonomy, and supply chain relationships data provider, and Truvalue Labs, an AI-driven ESG data vendor. Commenting on the Truvalue Labs deal in October 2020, FactSet's CEO Phil Snow said: “Truvalue Labs provides detailed and actionable ESG scores backed by a credible body of academics. ESG funds have seen record inflows this year and client demand for valuable, up-to-date ESG signals is high. Combining Truvalue Lab’s offering with FactSet’s content and technology suite will add value across our businesses.”

**AlphaSense’s Acquisition of Sentieo**

In May 2022, a market intelligence platform AlphaSense announced its acquisition of Sentieo, a research platform providing financial, social, news, web traffic, and app usage data. It follows the October 2021 acquisition of Mosaic, an expert call transcript library.

Jack Kokko, CEO of AlphaSense, said: “By welcoming Sentieo to the AlphaSense family, we’re adding over 1,000 customers, including over 800 institutional investment firms. We will soon be able to serve that combined customer base with even more best-in-breed functionality and content so that they have the insights they need to make the right decisions, including 1) Unmatched AI search, monitoring, extraction, and summarization capabilities; 2) Access to expanded content offerings for each company’s customers, including Stream’s transcript library; 3) Enhanced product workflows with Sentieo’s powerful technology and features, such as Table Explorer and advanced workflow tools that help users discover and work with key insights.”

**Tegus’ Acquisition of Canalyst**

In August 2022, research platform Tegus announced the acquisition of Canalyst, a financial model and company benchmarking data provider. Tegus Founders and Co-CEOs Michael Elnick and Thomas Elnick said: “Evaluating business and investment opportunities is an extremely inefficient process today. With our acquisition of Canalyst, we’ll be able to provide customers with a comprehensive view of the qualitative and quantitative data they need to uplevel and scale their investment decisions – all on a single platform.”
**Consumer Transaction Data**

There have been some developments in the consumer transaction space over the past year with a vendor having data disruption issues from a third-party channel partner.

This has stemmed upstream from Envestnet Yodlee, a 20-year-old financial data aggregator, whose partnership with a supplier of debit and credit card data appears to be in flux. As contracts are being renegotiated Yodlee has lost access to the data. As Yodlee is a supplier to several other channel partners, these partners have also seen disruption to their panels.

Vendors are now competing for market share and growing via acquisitions and partnerships. In an interview with Eagle Alpha, Jonathan Chin, Co-Founder, and Head of Data & Growth Strategy at Facteus commented on recent developments in the space: “The industry has seen a lot of groups selling the same or very similar products.

Anytime you see that consolidation could or should happen. It will be interesting to see how it plays out as many of the competitors have overlapping customers. Companies will have to make hard decisions about which products to support following an acquisition. There is no one-size-fits-all method of ingesting and delivering data products, so these companies will need to look at which ones make the most sense for their customers and ultimately, their profitability.”

In January 2022, Facteus announced the acquisition of 1010data’s Equity Intelligence business. Jonathan Chin explained the move: “Prior to the 1010data acquisition the Facteus data panel was heavily skewed toward debit cards and we had been actively onboarding new banks to add credit cards. The acquisition allowed us to accelerate the incorporation of credit cards into all Facteus products.

The addition of credit cards has enhanced our coverage and accuracy making Facteus transaction data products one of the top-performing alternative data sets in the market. We also benefitted from seeing their different approaches to normalization and productization. Every choice in product development has multiple right-answers, so seeing how and why they made their choices allowed us to hone in on our own value prop.”

It was also announced that Facteus and 1010data were planning to collaborate on retail and brand insights. Jonathan Chin highlighted that this would help Facteus introduce new products targeted towards fundamental investors:

These non-technical interfaces, we think, will play a big role in opening the fundamental investor market. In many ways, brands and fundamental investors have similar thought processes around uncovering the real drivers of market share. We see far more overlap in these deeper analyses than we do in just the topline, “call-the-quarter” work that transaction data is known for.”

There have been further developments in the consumer transaction data space as TransUnion acquired Verisk Financial Services in April 2022. Consumer Edge acquired Qentis, a European transaction data provider, in June 2022. And YipitData purchased Edison Trends in September 2022.

**Final Thoughts on M&A**

Early evidence shows that industry consolidation is happening across alternative dataset categories, including employment, web traffic, and consumer transaction. Further market moves will of course be dependent on demand from data buyers for various data categories and solutions and an overall recovery in global M&A which is at multi-year lows.
Contributors:

Dallán Ryan
Content & Advisory
Lead, Data Strategy

Mikheil Shengelia
Research Analyst, Data Strategy

Section 3

A Year in Thought Leadership: Trending Content for Data Strategy
Key topics discussed in 2022 include:

- The SEC’s exam risk alert and its impact on investment managers
- How a range of top hedge funds are implementing alternative data into their investment process
- Strategic insights and lessons learned from working with alternative data sources
- Sustainable investing, greenwashing, and ESG data
- Building internal data teams vs. outsourcing
- Alternative data growth for private markets and consultants
- Macro trends, inflation, and supply chain monitoring
- The rise of alternative data for crypto
- Regional data regulations: China, the US, Europe, Southeast Asia, Brazil, and India

Throughout the year we also received a range of requests to deliver analysis and insights on trending topics and datasets, including:

- Consumer transaction
- Data regulations in China
- Alternative data for fixed income analysis
- Supply chain data
- Web crawling
- Survey data
- Commodities
- Healthcare data
- Environmental and social data
- Sustainable Investing
- Employment
Legal and Compliance

Legal and compliance content received the most attention in 2022 as the SEC released its exam risk alert in April. This provided a catalyst for compliance teams to become more involved with the alternative data initiative. Our longstanding partners at Schulte, Roth & Zabel continued contributing to monthly workshops and content pieces with guest experts also making appearances. Outside of the SEC’s risk alerts, key themes we’ve seen growing interest in over the past year include greenwashing considerations, web scraping legal disputes, data privacy regulations, China’s changing regulatory landscape, and a much greater focus on compliance procedures.

Legal Workshops

In May, we hosted our best-attended client workshop discussing the SEC’s Exam Risk Alert where Ben Kozinn, Jessica Margolies, and Kelly Koscuizska from Schulte, Roth & Zabel discussed the implications of the SEC’s announcement. One area of concern for examiners was how firms handled potential MNPI derived from alternative data, defined as information about a company’s performance found outside of financial statements, company filings, and press releases. The Schulte, Roth & Zabel team highlighted several recommendations:

1. “You need to have robust policies and procedures if you will be using alternative data. The SEC is focused on asset managers and how they use alternative data sources, and it is not enough to only have policies and procedures related broadly to research. They want to see something specific regarding the use of alternative data.”

2. “Use a risk-based approach during the diligence process. There is no one-size-fits-all model. You need to have a diligence process for different types of vendors. The SEC will pay attention to how well you demonstrate the risks that specific vendors bring.”

3. "Consistent process of onboarding and periodic reviews of vendors. In addition to onboarding a new vendor, regular diligence is critical. New datasets from existing vendors should trigger a reassessment of said vendor - a recent conversation, perhaps a reaffirmation of the DDQ. Any negative news that you hear about the vendor might trigger additional diligence. Legacy vendors that were on boarded before you implemented new procedures should also go through the assessment.”

In August, we were also joined by a former SEC examiner Peter Greene from Schulte, Roth & Zabel Igor Rozenblit for the MNPI Compliance Considerations for Investment Advisors workshop who provided an overview of the SEC’s approach to cases involving MNPI and discussed compliance procedures.

In April, for the Alternative Data Compliance Process workshop, we were joined by the Schulte, Roth & Zabel team along with Michael Abbriano of ACA Compliance Group to discuss the steps in the alternative data compliance process for buy-side firms. This conversation complemented our report on the topic with discussions focused on team structure, sourcing and data identification, web scraping, trial agreements, standardization of documentation, onboarding, monitoring.

Readership

Based on client engagement, our buyside legal and compliance survey was the most-read legal article of the year. Our ongoing coverage of the hiQ vs. LinkedIn case received the second-highest readership for the year. We also updated our Data Privacy and Alternative Data report specifying regulatory updates and recent developments surrounding the mobile app, web scraping, and web traffic data.
Use Case Insights

In 2022, we covered a wide range of dataset categories, including event data, supply chain data, ESG, social media and sentiment analysis, insurance data, alternative data for commodities, insurance data, macro data and news, employment data, and survey data. We will continue to deliver content across our new and improved dataset category taxonomy throughout 2023.

Workshops

Our most attended alpha workshops in 2022 were on social media and web data closely followed by the webinar on employment data. In our September client workshop on Social Media and Web Data for Tracking Online Trends and Sentiment, we took a deep dive into social media and web data using social listening, sentiment, and signals to track trends and brands online. Social media data and sentiment are one of the most popular alternative data categories and are used by asset managers in competitive analysis, to track conversations around brands and engagement, to understand consumer trends, and to analyze broader industry and sector shifts.

Our April Employment Data Workshop explored use cases and trends including the relationship between employee turnover and firm performance, the rise of the chief sustainability officer, and employee perception analysis of DE&I in the Fortune 100.

Readership

Our most popular use case content in 2022 was the article on Alternative Data for Crypto Markets where we discussed the most established alternative data providers offering data to cryptocurrency investors.

From on-chain analysis comprising a record of every step linked to the transaction and the modification of the blockchain following it to off-chain study consisting of steps outside of the blockchain, the breadth of information available to investors is far superior to that of traditional equities.

The spotlight report on Working with Survey Data also received a lot of attention this year. A report discussing Working with Survey Data also received a lot of attention this year. We explored market research and surveys as an alternative data source, highlighting leading vendors in the space and providing expert perspectives on themes associated with surveys for financial services and corporates.

Similarly, a report on Monitoring Employment Trends with Alternative Data has also received interest as employment data was the 4th most popular category over the past 12 months. We saw a 40% YoY increase in the number of employment datasets onboarded. The report included an overview of the leading vendors, deep dives & use cases, and legal considerations.

This report is complemented by our survey data workshop where we were joined by Professor Anatoli Colicev from Bocconi University, Michael Mayhew, Chairman of Integrity Research, and Matei Zatreanu, CEO of System2.

“The spending on market research has rapidly increased from $33.15bn in 2009 to $73.38bn in 2019. Corporations were found to be the largest users of surveys for business planning purposes, while only 4% of total spending on surveys came from financial services.” – Michael Mayhew, Integrity Research
Strategic Insights

Our report on Monitoring Employment Trends with Alternative Data has also received interest as employment data was the 4th most popular category over the past 12 months. We saw a 40% YoY increase in the number of employment datasets onboarded. The report included an overview of the leading vendors, deep dives & use cases, and legal considerations.

The Strategy Centre was introduced in 2021 serving as a one-stop-shop of workshops, video content, and articles to help firms establish best-practice alternative data strategies and stay ahead of the competition. This stream has grown in popularity inviting experts to discuss their experience in client workshops and reporting on industry trends through a strategic lens.

Workshops

Our most popular Strategy workshop was in September when we were joined by Former McKinsey Chief Data Officer Mohammed Aaser&former McKinsey Chief Data Officer Mohammed Aaser who discussed his illustrious career and experience working with alternative data. While at McKinsey, Mohammed launched the data innovation team, partnering with divisions to incubate and design new data assets, resulting in the development of 40+ data assets and partnerships with 50+ data providers.

He also developed and implemented a firm-wide data catalog, indexing 1,500+ proprietary, and external datasets, building a data steward community (1,000+ data experts), and leading change management efforts across the organization, with 15,000+ users supporting thousands of client engagements.

Mohammed highlighted multiple challenges when working with data vendors such as consumer panels and how the sourcing approach is not clear as many vendors use a lot of marketing jargon without explaining exactly how they get the data. Legal and compliance considerations and licensing issues are also high on the list. Mohammed explained: “There are frictions across the entire value chain, so you end up spending 80% of your time understanding the data space, investigating it, onboarding, and so on, and then you only spend 20% of the time actually using it.”

Readership

On the engagement side, we found that our analyst note on Consumer Transaction Data Disruptions saw the highest readership over the past 12 months. Envestnet Yodlee’s partnership with a supplier of debit and credit card data appeared to be in flux resulting in Yodlee losing access to the data. As Yodlee is a supplier to several other channel partners, these partners have also seen disruption to their panels. We highlighted recent developments in the space across the category and impacting different vendors and what is important to understand from a buyer’s perspective.

As part of our strategic push, we published two key strategic pieces that received a lot of attention. Our report on Data Sourcing Best Practices pulled together publicly available data and insights from industry experts and Eagle Alpha’s proprietary content to discuss the data sourcing team and workflow; challenges and solutions; legal and compliance considerations; buyside case studies. While our report on Outsourcing Considerations highlighted a broader trend we see in the alternative data industry by firms at all stages of the data journey, even the most data-sophisticated firms.
Technical Insights

This section of our Data Strategy portal is devoted to technical topics such as data science and engineering lessons learned, data hackathons, nowcasting and NLP techniques, etc. Below we highlight the most engaging pieces from 2022.

Workshops

Our workshop on Eagle Alpha’s Data Style Guide was prepared following client requests with our CTO Thomas Combes taking a deep dive into best practices for data hygiene. The workshop was complemented by the report highlighting file type, data type, publishing, and general data best practices. Thomas outlined there has been niche use for formats like XML, and he has never heard of anyone using Avro due to its many pitfalls. Additionally, the case for using CSV is very minimal. While CSV is the de facto standard for tabular data, it has significant limitations.

Thomas recommended using Parquet. In his experience, vendors usually target CSVs as the end goal assuming end-users will open files in excel. Thomas believes that if there are 1000s of files in CSV format, the end-user will often only open one of them for inspection. If completing a delta-style publication is a method of publishing ongoing data products where only new data points are issued, the assumption is that it should include only new pieces of data, recent events, and new rows. However, datasets like a significant consumer transaction dataset can contain unusual structures that you must also consider.

In February 2022, we hosted the INTERACT 2.0 virtual data conference. As part of this event, Seth Leonard of System2 led a conversation exploring nowcasting about inflation, lockdowns, and the consumer. The panelists discussed how transaction data has helped navigate the volatile macro and micro environments over the last two years and trends towards higher frequency inputs into macro models.

Readership

Our CTO Thomas Combes authored an article titled “The View from Lakeshore” where he defined a data lake as underpinned by a cloud-based object storage system that has API-based access methods (Amazon S3 being the most recognizable example). A sample of the full article is presented below.

One of the best and worst aspects of a Data Lake is that there are very few restrictions on the data structure. Whereas a Data Warehouse usually has Databases, Schemas, Tables, etc., the object storage systems of Data Lakes have only very base-level organizational concepts, mainly files, and folders.

While this might sound elegantly simple, having almost no restrictions on file formats, directory organization, or file naming conventions turns into a detriment when having to share your Data Lake with your colleagues, let alone having to share it with a 3rd party that has different ideas on how best to organize and present data.

*We found ourselves needing to know with a high degree of accuracy what we received access to and to build up that awareness very rapidly to fulfil these delivery requirements in tight timeframes. We tried a few 3rd party tools to “scan the data”, each of which broke, sometimes silently, from the assortment of edge cases we were experiencing, resulting in erroneous information on what we were given access to, and leading to costly and manual rework.*
Public Content

It is important to note that Eagle Alpha’s Data Strategy team also produced content for a wider public in the past 12 months, including webinars, podcasts, and blog posts. The most notable pieces of content are presented here.

• In March 2022, we hosted an ESG data-focused webinar alongside Gerald Garvey, Head of Sustainability Research at BlackRock, and Adam Gustafsson, a Quantitative Researcher at UBS. The subsequent report with presentation summaries is available on our public-facing website.

• In April 2022, Bain & Co., Ernst & Young, West Monroe, Asgard Data, and Strategy Titan joined us for a deep dive webinar into external data for consultants and the success they have seen. The subsequent report showcased the art of what is possible with external data through the lens of consultants working with corporate and private equity to implement data-driven decision-making.

• In June 2022, we updated funds and investors on the latest data and use cases relating to alternative data for macro applications. We discussed a new market regime for CTA and Macro investors, with comments from Bank of America, and macro strategies incorporating new data sources.

We also published a range of blogs including: Supply chain disruptions, Alternative data in emerging markets, The rise in data for crypto investors, Tracking employment trends, Alternative data for private markets, The data sourcing overview, and The dataset evaluation challenge which we co-wrote with our strategic partners, Exabel.

Conferences in 2022

We hosted several conferences this year with highlights of each included on our Data Strategy portal.

• In February 2022, we hosted our INTERACT 2.0 virtual conference. Key presentations included The Shanghai Data Exchange by Trivium China; New Applications of External Data Within the Insurance Sector by McKinsey; Cloud, Data Warehousing, and a Deep-Dive into Data Science Applications by Man Group.

• In April 2022, we hosted our UNBOUND conference, the first in-person event of the year in Midtown New York. Key presentations included Double Bottom Line Framework by BlackRock, and our fifth Data Hackathon with System2, Teza Technologies, and Crabel Capital. The research question was: “The Endemic & War – Winners and Losers in Hospitality, Travel & Leisure”. The winning team used a combination of earning calls data, ESG data, and employment data to develop a portfolio with a Sharpe Ratio of 1.67.

• In June 2022, we held our External Data for Private Markets conference in Boston. Key presentations included Alternative Data Use Cases for Private Equity by Bain & Co, The Rise of Data Talent in Private Markets by Revelio Labs, and panel sessions included discussions of alternative data applications for deal sourcing, due diligence, and valuations.

• In October 2022, we hosted our first conference in London since 2019 where notable topics of discussion included greenwashing, consumer transaction data disruptions, macro trends, private markets data adoption, the changing role of the compliance function, and combining data with fundamental research.
As mentioned throughout the report there has been a range of important instances in 2022 that have shaped both the data world and everything that surrounds it.

This section is a collection of significant contributions from industry experts who provide insights into timely topics such as inflation, crypto, supply chain and trade, employment, and more.
Inflation and Financial Markets

Contribution from Turnleaf Analytics

Forecasting Inflation in a Big Data world

To account for the nonlinear relationships between many of the variables used in the Big Data world, we can use machine learning models which can capture these dynamics, which cannot be captured as well by simpler models such as OLS. At the same time, we want to use models that are still fairly tractable and understandable. Turnleaf Analytics’ approach has been to embrace the latter way to forecast inflation using machine learning. They also use alternative data to augment more traditional datasets such as those containing macroeconomic and market data.

Turnleaf Analytics’ country universe consists of 28 countries across both emerging and developing markets. Every month after each inflation print, they update the forecasts for each country covered, for horizons from 1 month out to 12 months. For each time horizon, they fit the model. Hence, for each country, they have 12 models. The rationale is that the inflation pass-through can be quite different depending on your forecasting horizon.

In the short end, you might find that yields are correlated to inflation because rising inflation is often accompanied by central bank hikes. By contrast, in the back end, the correlation becomes negative, as the rates hikes pass through into the broader economy.

Since May 2022, Turnleaf Analytics has made nearly 200 live forecasts which they have been able to compare against actual inflation releases. In this live environment, their model has beaten the benchmark forecast (drawn from official sources such as the central bank) over 60% of the time. The mean absolute error is also 20-30% lower than the benchmark.

Turnleaf Analytics’ Latest US Inflation Forecast

In Figure 4, they show the latest inflation forecast for the US. In the short end, the model expects less inflation than the market. As we go further out, the situation is more mixed. However, both the model and the benchmark suggest inflation in the USA has peaked. They note, however, the picture does vary significantly across both developed and emerging markets.

Figure 4 – Turnleaf Analytics’ Model forecast for Inflation in the US Versus the Benchmark (Source: Turnleaf Analytics)
Creating a Trading Rule for US Treasuries Using Inflation Forecasts

There are many use cases for inflation data in markets, ranging from trading inflation swaps and bonds to equity indices. Turnleaf Analytics recently wrote a paper that used our long-term inflation forecast as input into a trading strategy for US Treasuries and TIPS ETFs, which is summarized here.

They created a metric based on the long-term inflation forecast against the market pricing for inflation. When this metric is positive, it means the model is expecting more inflation than the market and this triggers a short bond position. Otherwise, the model is long bonds, given their model expects less inflation from the market.

In Figure 5, they plot a stylized example of this metric, and alongside it they also plot UST 10Y yields. They note, that at least in their stylized example, the inflation metric was a leading indicator for the decline in yields seen in 2019, and also flipped to positive, before yields began to move higher.

In the short end, their model expects less inflation than the market. As we go further out, the situation is more mixed. However, both the model and the benchmark suggest inflation in the USA has peaked.

In the short end, their model expects less inflation than the market. As we go further out, the situation is more mixed. However, both the model and the benchmark suggest inflation in the USA has peaked.

In Figure 6, the active inflation-based trading rule outperforms long only for bond ETFs (Source: Turnleaf Analytics)

Conclusion

Forecasting inflation is a difficult problem to solve. However, using newer approaches such as machine learning and using alternative datasets as part of the process can help to increase forecast accuracy. Turnleaf Analytics have also shown how to use their model-based inflation forecasts in the context of trading rules for US Treasuries and TIPS, which have historically outperformed long-only positions.

Whilst, past performance is not indicative of future performance, in a historical sample going back to 2017, their active trading rule, using the above inflation metric as an input, outperformed long-only strategies for both US Treasuries and TIPS instruments (see Figure 6). In particular, risk-adjusted returns were -0.37 in the historical sample for long US Treasuries ETFs, the active strategy had risk-adjusted returns of 1.17.
The Volatility of the Crypto Markets
Contributions from Kaiko and Coin Metrics

The FTX Collapse – How Cryptocurrency Liquidity Data Can Improve Risk Management

By Kaiko

A neglect of proper risk management and a poor understanding of liquidity caused some of the largest crypto firms to become insolvent throughout the year. More than $14 billion of FTX’s balance sheet was held in a cryptocurrency called FTT, created by the exchange. The problem with FTT was that the token was relatively illiquid, with very little trading activity and buy-side demand. Just a small trigger caused a wave of selling pressure that the market could not support.

Using Kaiko’s cryptocurrency liquidity data, we can see that market depth for FTT was just $8 million dollars before the collapse, a strikingly low number considering the token had a market capitalization of more than $3 billion.

As fear gripped the market and sellers rushed to the exit, FTT’s price quickly dropped as bid-side liquidity evaporated and market makers removed orders. As of November 25th, the market depth for FTT is just $50k across all exchanges.

FTT investors are now left holding a token that can’t be sold due to poor liquidity. Had FTX and other FTT investors understood the token’s low liquidity and possible risks involved beforehand, they could have set up proper risk controls and avoided the collapse.

Understanding liquidity is thus a vital component in a robust risk management framework. The valuation of a fund’s balance sheet is only as strong as its ability to efficiently liquidate its holdings. Liquidity metrics, combined with traditional risk metrics such as VaR or Expected Shortfall, will enable investors to better survive the next bear market.

![FTT Market Depth](source: Kaiko)

**Figure 7 – FTT Market Depth (Source: Kaiko)**
A Front-Row Seat to FTT’s Collapse

By Coin Metrics

CoinMetric’s analysis of Binance’s markets renders a dramatic scene of the events leading to the catastrophic devaluation of FTT. Binance (which is the largest crypto exchange in the world by quite a margin) is of particular interest due to its pre-eminence, the extent of its derivative products, and (critically) the significance of Binance’s own FTT market.

Coin Metrics ventured to compare the price discovery mechanism operating on Binance relative to other FTT trading venues and built an orderbook heatmap—leveraging Coin Metric’s Market Data Feed (MDF)—to perform a complete analysis of Binance’s FTT market, starting on November 1st, until the complete collapse of FTT by November 11th. They especially wanted to understand the market in the aftermath of the CoinDesk article, published on November 2nd, which revealed that much of FTX’s balance sheet was FTT.

Shortly after the CoinDesk article was published, a clear price support formed on Binance’s orderbook at $23.50 per FTT. Their hypothesis is that Alameda was signaling to the market their willingness to defend FTT. Two days later, on November 5th, large sell orders shifted the price trend. This was a day before CZ (the CEO of Binance) announced the unwind of their FTT position, and the shift happened on Binance first.

On November 6th, the announcement arrived, and with it came a new wave of price support orders at $20 and $10 per FTT. They hypothesize that these were also placed by Alameda. There are clear attempts at reviving FTT’s price throughout the day, especially after a tweet by Alameda’s CEO offering to buy all FTT owned by Binance. These attempts ultimately failed as sell orders kept coming in. The collapse arrived early on November 8th, with the announcement that FTX was pausing withdrawals.

This was likely not the first time Alameda was buying FTT. The token’s significance as FTX’s largest position, its use as collateral for loans, as well how its price functioned as a barometer of confidence in FTX makes FTT central to the story of FTX and Alameda. It is almost certain that FTX user funds were used to prop up the price of FTT, especially toward the end. We will hopefully find out more in the following months.
S&P Global’s most recent data for mainland China show that after a positive trend in year-on-year growth for exports in June and July 2022, a significant slowdown in mainland China exports can be observed in August (7.0% growth y/y) and September (5.6% y/y). When it comes to imports, August and September 2022 already brought negative growth rates amounting to -0.2% y/y and -0.4% y/y, respectively.

Looking at the sectoral breakdown, a few groups of goods for which Chinese exports plummeted in September 2022 can be named, like Optical, Photographic, and Medical Apparatus, Toys, Iron and Steel and Textiles, Apparel, and Clothing marking from 10% to nearly 30% drop year-on-year.

Using GTAS, recent trade patterns for the main potential beneficiaries of China’s slowdown can be analyzed:

1. Indonesia is showing the opposite trend to China in the latest months with nominal value year-on-year growth rates for exports to the United States, Netherlands, and United Kingdom above 15% in August 2022.

2. The Philippines is one of the main producers of nickel – a mineral used for producing electronic devices and batteries; United States imports of electronic integrated circuits and micro assemblies from the Philippines is increasing in recent months.

3. Vietnam – some investors moving away from China perceived as the most beneficial location; the GTAS platform will show you that in the third quarter of 2022, Vietnam’s exports continued to perform very well, remaining more resilient to supply chain disruptions than earlier. The compound annual growth rate (CAGR) for Vietnamese exports for 2022-2025 is estimated at 4.2%, according to the GTAS Forecasting model.

4. Bangladesh – GTAS Forecasting projects total exports from the country to increase by 9.3% year-on-year in 2022 with the majority of this growth driven by apparel and clothing accessories exports (estimated to reach 48 billion USD in 2022). The country’s apparel and clothing accessories exports increased nearly 25% year-over-year in 2021 and are expected to grow 12% in 2022.
In May 2022, the trend toward more layoffs—particularly in Tech—has increased since the initial waves of job market volatility. On December 1st, global outplacement firm Challenger, Gray & Christmas published an estimate of job cuts for November that was up 127% over October and the highest since January 2021. Using data from https://layoffs.fyi/, LinkUp set out to see if job listings still signal these layoffs.

They now have a sample size of 546 companies that have coverage before their layoff announcement. The average percentage change of listings prior to the layoff announcement date remained consistent with what was observed in their first post: the 30-day change was at -27%, the 90-day at -45%, 180-day at -54%. The early warning signals are still present for companies approaching layoffs.

**Robinhood**

A slowing in retail trading generally, and a downturn in the crypto market specifically, likely led to Robinhood having to cut roughly 23% of its workforce. Job listings decreased sharply before both instances of their layoff announcements. Somewhat of a turnaround does seem to be in place, however, as lately their listings have been on the rise.

**Carvana**

A cooling used car market largely attributed to higher rates on auto loans, and a balance sheet weighed down by debt, created a situation where Carvana had to implement multiple layoffs in an attempt to return to profitability. In both cases, listings fell sharply before the layoff announcements.

**Microsoft**

The tech behemoth’s size did not mean it was immune to layoffs; Microsoft cited business priorities as the driver. In both instances, LinkUp-collected job listings had decreased sharply. Now it does appear those listings are showing signs of stabilizing.

---

In May 2022, the trend toward more layoffs—particularly in Tech—has increased since the initial waves of job market volatility. On December 1st, global outplacement firm Challenger, Gray & Christmas published an estimate of job cuts for November that was up 127% over October and the highest since January 2021. Using data from https://layoffs.fyi/, LinkUp set out to see if job listings still signal these layoffs.

They now have a sample size of 546 companies that have coverage before their layoff announcement. The average percentage change of listings prior to the layoff announcement date remained consistent with what was observed in their first post: the 30-day change was at -27%, the 90-day at -45%, 180-day at -54%. The early warning signals are still present for companies approaching layoffs.

**Robinhood**

A slowing in retail trading generally, and a downturn in the crypto market specifically, likely led to Robinhood having to cut roughly 23% of its workforce. Job listings decreased sharply before both instances of their layoff announcements. Somewhat of a turnaround does seem to be in place, however, as lately their listings have been on the rise.

**Carvana**

A cooling used car market largely attributed to higher rates on auto loans, and a balance sheet weighed down by debt, created a situation where Carvana had to implement multiple layoffs in an attempt to return to profitability. In both cases, listings fell sharply before the layoff announcements.

**Microsoft**

The tech behemoth’s size did not mean it was immune to layoffs; Microsoft cited business priorities as the driver. In both instances, LinkUp-collected job listings had decreased sharply. Now it does appear those listings are showing signs of stabilizing.
Wayfair

Similar to others, Wayfair laid off around 5% of its workforce citing overestimates of growth that resulted in a team too large for the current environment. Prior to this layoff announcement, Wayfair’s job listings had fallen sharply.

![Figure 15 – Job Listings Data for Wayfair (Source: LinkUp)](image)

Amazon

Amazon made headlines when they announced they were laying off 10,000 employees. LinkUp-collected listings hinted that trouble was on the horizon as listings fell sharply. It is worth noting that their Amazon scrape did break during this time as well, but the downward trend was present before this break.

![Figure 16 – Job Listings Data for Amazon (Source: LinkUp)](image)

Splunk

The last time Splunk announced layoffs their listings had a similar drawdown. Furthering the case for possible layoffs in the future, new management is taking over and is likely to implement material changes. Additionally, activist investors, known for pushing operational efficiency, have taken a large stake in Splunk and may advocate for improved operating margins. On the other hand, it is worth noting LinkUp has seen large drawdowns in Splunk’s hiring before with no layoffs.

![Figure 17 – Job Listings Data for Splunk (Source: LinkUp)](image)

TripAdvisor

TripAdvisor is coming off a quarter where they saw strong revenue growth. At the same time, selling and marketing expenses grew significantly, decreasing profitability and setting the backdrop for a possible reduction of headcount. LinkUp can already see that TripAdvisor is slowing its anticipated hiring pace.

![Figure 18 – Job Listings Data for Trip Advisor (Source: LinkUp)](image)
More and more data specialists have started arriving at small and medium-sized private equity firms, and employment data can be helpful in this type of industrial analysis. Revelio Labs aims to create the world’s first universal HR database by absorbing and standardizing hundreds of millions of public employment records. It is then possible to analyze workforce dynamics and trends for any given company.

For this research question, Revelio Labs first needed to specify the universe of private equity firms. Revelio Labs prepared a list of around 1,400 companies identified as such. Using detailed role taxonomy, their team then specified roles based on activities to compare apples to apples. The results show that private equity firms’ data analytics-related roles have doubled over the last decade.

All other roles are growing, too, so we can conclude that the whole private equity industry has been on the rise for the past decade. However, the growth of data-related roles has outpaced everything else. Leaders in talent for data analytics were also identified, although the results were not normalized. Therefore, we can expect that larger companies will have more data talent.

The role taxonomy that Revelio Labs uses is based on activities rather than titles. The company uses NLP and pre-trained word models to cluster roles into various activity groups. Lisa Simon, a Senior Economist at Revelio Labs noted: “It is really what someone does, not what they say they do because titles mean very different things across different companies.” It is also possible to dig deeper and analyze the skill set of a given talent pool.
When it comes to analyzing earnings calls, many NLP providers have attempted to capture the sentiment from a call, often comparing the sentiment from management-prepared statements (Management Discussions) versus the sentiment from Questions and Answers from analysts. While other firms derive sentiment from past returns, sentiment dictionaries, or pre-trained models, Alexandria believes it is the analyst themselves that are the best source for sentiment training data.

The company has worked with buy-side and sell-side analysts to label sentences in news, calls, social media, and more over the last 10+ years. The goal is analyst automation to solve a Big Data problem. The thinking goes, if an analyst had the time to read every document on a company, this would be the best resource for the sentiment. Since an analyst can’t read even 1% of the information available, the next best option is to use artificial intelligence to model their theme and sentiment selections.

The proof is in the returns. Alexandria’s NLP, trained by analysts, outperforms dictionary approaches and BERT-based models, demonstrating that analysts really are the best starting point for machine learning. Alexandria applies this same approach to news, social media, and long-form documents on various assets from companies to macroeconomics and commodity to crypto.

Figure 21 – NLP Capabilities of Alexandria Technologies (Source: Alexandria Technologies)
Section 5

The Year in Market-Driven Product Development
A core part of Eagle Alpha’s culture is constantly improving our products to meet and solve the challenges our customers face every day. The feedback we received from buyers and data providers meant that the roadmap to their challenges was ever-changing based on the evolving marketplace.

For example, implementing a more entrepreneurial feature development framework in 2022 improved the focus to be on addressing complex problems on both sides of the market. Such as connecting production datasets to our platform to generate rich and informative data profiles. This solved being more responsive to emerging customer needs such as the requirement for compliance teams to monitor and track their dataset due diligence activities.

The feedback we received from clients has been extremely valuable and we want to take this time to thank our clients for the honest feedback over the past year that has helped shape our roadmap and deliver problem-solving solutions directly to you. Here is an overview of the features we released this year, in response to the industry challenges.

**Updated Data Taxonomy**

In 2016, we were the first company to launch an alternative data taxonomy. Our honeycombed 26-category alternative data taxonomy was ground-breaking when it first launched and has been adopted by leading global data buyers and data vendors to categorize types of datasets in the alternative data market.

This original taxonomy has not grown in size since it was launched over 7 years ago, but the industry has and we believed that it was time to evolve and update our taxonomy to provide a more robust and granular format for both buyers and vendors in the market and their needs. Figure 2 below is our classic taxonomy used by the alternative data industry.

We originally launched with 24 key categories which then expanded to 26 and over the past year it became obvious that this taxonomy was no longer granular enough and there was a lot of overlap between categories. One major example is Web scraped data. There can be multiple types of web scraped data like pricing, social media, sentiment, news NLP, reviews & ratings, and so on. It is obvious now that Web Scraping should be classified as a technique and not a data category so in our new taxonomy web scraping does not feature as a category.

We are proud to announce that our new data taxonomy is a comprehensive framework to classify data types by what type of data a source offers. It provides a granular view for data buyers and data vendors by classifying data using a bottom-up approach according to the data that a data vendor provides.

![Figure 22: Eagle Alpha’s Traditional 26-Category Alternative Data Taxonomy Launched in 2016](image)
The taxonomy has a new hierarchical structure with 16 main categories that break out into 56 sub-categories.

This updated taxonomy has enabled us to capture newer data types like Crypto data whereas in the previous taxonomy it was difficult to classify these types of datasets. The taxonomy will allow data buyers to better classify data and help data vendors understand what type of data to tag their datasets to. The below graphic reveals our new data taxonomy broken out by category and sub-category.

Compliance Workflows

An emerging theme at the end of 2021 that then ramped up in 2022 was the need for better processes and tools to perform due diligence on data products in all phases of the data procurement lifecycle. Over the past year, the SEC has increased its scrutiny and risk profile of alternative data requiring buyers to maintain an audit trail of their due diligence when seeking to acquire and work with alternative data by characterizing the dataset, the data vendor, and the alternative data market.

Figure 23: Eagle Alpha’s New and Improved Alternative Data Taxonomy
Compliance Workflows Cont’d

To address these challenges, we started by compiling the largest and most up-to-date database of vendor-supplied due diligence questionnaires (DDQs), achieved through a simple web-based form on the vendor side. Instead of vendors filling out specific due diligence documents for each buyer and having to save their responses in a static PDF format.

Now vendors can manage and update their responses to an FISD-standard Due Diligence Questionnaire all via a web-based form right on their data profile. For data buyers needing to perform ongoing monitoring of data vendors, we incorporated the full lineage of questionnaire responses onto the data profile. We also released a feature to allow users to set up email notifications related to any changes in DDQ responses of any dataset that is important to them, increasing the ability of compliance teams to respond to changes in the legal and compliance posture of a data product in a timely manner.

Connected Data Profiles

In 2022, we set out to become a trusted market validator for data products found on our platform. While Eagle Alpha historically has been relied upon for “mapping the universe” of alternative data through metadata gathering, this year increasingly data buyers wanted to vet data products before making a purchase decision.

In addition to utilizing our trial facilitation service which adds a layer of standardization to the buyer’s data assessment workflow, data buyers can now request that data vendors link their data products to their profiles. The insights and validation that connecting a live production data product to our platform delivery began to be seen as a must-have for buyers assessing data products. Having a trusted 3rd party validate baseline characteristics of delivery consistency, data structure, and coverage history give a degree of confidence for data buyers unfamiliar with a data product. It also helps new vendors build up their credentials as a trusted counterparty quickly, leveling the playing field between market incumbents with established reputations and disruptors with new and differentiated products.

Benchmarking Datasets to Peers

Selecting the right type of data to answer a research question is a challenge by itself, but when you are met with a number of similar datasets from different vendors, being able to compare the attributes of different datasets becomes increasingly important. We sought to address this challenge with the release of two complimentary features: live-computed peer benchmarking and the display of related datasets to any dataset, both available directly on the data profile to Eagle Alpha buyer clients.

The benchmarking to peers is performed for 9 key dataset metrics, such as the length of dataset history or number of vendor employees, compared to the peer group of approximately 30 most similar datasets.

That peer group is derived from a dynamic measure of dataset profile similarity including fields such as the dataset description, data category, and use case descriptions to name a few.

These two complementary features allow data buyers to become more informed shoppers when evaluating the market for a type of data product, and feel confident that they have assessed all avenues for addressing their data requirements, including the often-overlooked ones.
Conference Meeting Recommendations

In 2022, our conferences returned to in-person, and we quickly took the lessons learned from hosting many virtual conferences and applied them to the in-person setting. This resulted in our platform becoming a primary part of the experience of attending our conferences.

Based on attendee feedback, we focused on improving the platform experience by releasing simpler password-less login, adding connected data elements to data profiles of attending vendors, and reducing the number of clicks needed to book a meeting by releasing a vendor-level summary page with direct booking functionality.

Often buyer and vendor attendees at an alternative data conference are not aware of which vendors and buyers would be suitable for each other to meet based on data requirements and characteristics.

For our London conference in October 2022, not only did we roll out an email-based meeting recommendation system personalized for each data buyer attendee, we also released a feature to allow vendors to request firms they would like to meet with from the list of firms registered to attend that conference. These vendor requests were sent to data buyers prior to the day of the conference, resulting in a substantial number of matched meetings taking place where the preferences of either the buyer or the vendor (or both) had been fulfilled.

These features combined to result in positive feedback from many attendees about both how easy and how effective attending an Eagle Alpha conference was for them, and we plan to continuously improve in this area for our upcoming conferences, the next being in January 2023 in New York.

Collection Notifications

Data Collections were a prominent new platform feature we introduced in 2021 that we improved upon in 2022. This year we introduced the first notification types to collections. Users can now select to receive notifications when a vendor updates a DDQ for any dataset in a collection. Users can select the frequency that they want to be updated (daily, weekly, or monthly), as well as opt-in to be notified only about response changes to specific DDQ questions they select.

The notifications are sent via email, outlining the collection, dataset profile, and DDQ field that was updated, as well as the vendor’s previous response. Our compliance team users gave us feedback that having an audit trail of notifications about pertinent information from a vendor changing was important to them to meet emerging regulatory requirements, and we plan to continue building out solutions to these requirements in 2023 by improving the monitoring, tracked communications, and workflow management capabilities of the platform.

Self-Service Data Connection Process

As mentioned above, in 2022 we began to see data buyers come to us requesting to see the Eagle Alpha validation that comes from a vendor connecting their dataset to their data profile on our platform. We focused on making this process as streamlined as possible, achieving the possibility of a completely self-service setup of a data connection if the data vendor was publishing their dataset to an AWS S3 bucket already.

This means that a data vendor that wants the validation of a connected data profile can achieve this, through a process that can be completed on platform in under 20 minutes, all without needing manual intervention from Eagle Alpha staff.
This process can end up being as simple as three steps: 1) Supply the AWS S3 Bucket Name and Prefix pointing to where the data is published to, 2) Apply a custom-built bucket permission policy to allow Eagle Alpha read access to the bucket, 3) Verify the connection is established by reviewing the data dictionary and samples we generated automatically by scanning the bucket contents.

Removed is the need to constantly update and upload a new data dictionary file and sample data whenever column names change or new columns are added, as the connected data profile automatically evolves as the data does.

**Updated Data Vendor Profile & Homepage Design**

We addressed vendor feedback in 2022 that maintaining and creating dataset profiles was not as intuitive as it could be by overhauling the user experience. Now vendors can update their live profiles in real-time, and importantly get a view of what their profile looks like to buyers while they are still in the editing process. While a vendor is filling out a profile, the updating profile completeness score makes them aware of how their profile compares to their peers in terms of completeness of information, helping to focus the contribution on making their profile rank highly in the market by offering a transparent and detailed view of their offering.

**2023 Product Plans**

Building on our momentum in 2022, the outlook for product development in 2023 is very exciting.

The goal for our buyer platform in the year to come is to build more mission-critical functionality that our customers have been asking for. When a data buyer client is trying to decide whether to commit to purchasing a data product, we want to be there to build confidence and provide clarity in a way that only a trusted third-party with access to, and knowledge of, the actual data product can bring. We want to make it easy to compile all of the work that has gone into assessing different aspects of a data product into a coherent, reviewable form so purchase and renewal decisions are made with full stakeholder visibility.

For the data vendor platform, our goal in 2023 is to provide increased insight into the market dynamics for a data product and hone our suitability identification of data buyers with new sources of information such as hiring trends of data specialists at these institutions. We want to make portable the data profiles that vendors contribute so much time and effort to complete, and have them be useful far beyond the display of relevant metadata about the product. Making trials and subscriptions, from due diligence requests to data delivery, less cumbersome for data vendors is a key area of focus in our vendor product development in 2023.

As a thank you to all of our clients who gave us feedback in 2022, we want to acknowledge that it has been extremely helpful in getting our products to where they are today, and we will continue to listen to our clients about their problems throughout next year, and work tirelessly to build products to address these challenges.
So, what are our key predictions for 2023:

1 Compliance scrutiny of data provenance is going to increase. Compliance teams will look for an increasing level of transparency in the data supply chain, and the regulator will look for more time stamping and audit trails that capture this. Some products and categories will find it increasingly difficult to reach compliance thresholds.

2 Private equity usage of data for due diligence and portfolio companies will continue to grow and data expertise within private markets firms will continue to expand and ensure they are more comfortable with ‘raw data’ like their public market relatives.

3 Vendors and buyers within the private equity vertical will consider fresh approaches to commercials, data extraction, and delivery exercises to allow alternative data to integrate better into private equity exercises.

4 Compliance officers will utilize more tools, platform functionality and time stamping that go beyond static document capture and 3rd party opinions to align with regulatory needs.

5 The sell-side banks will utilize more alternative data within their research. The gap between buyside and sell-side approaches to company analysis has grown. The former has become more data and expert network driven this gap needs to close.

One of the most important developments at Eagle Alpha this year has been the evolution of our platform in relation to data product insights and compliance. The platform features became more relevant due to the SEC Risk Alert in April which provided guidelines to compliance teams at investment advisors where their firms were utilizing alternative data as it relates to material non-public information (MNPI). The SEC Exams Division said that compliance teams needed to improve their ‘re-profiling’ and monitoring of risks relating to dataset compliance during the product lifecycle.
About Us

Established in 2012, Eagle Alpha is the pioneer in connecting the universe of external and alternative data. We are the leading data aggregation platform with supporting advisory services for data buyers and data vendors.

First adopted by alpha-seeking hedge funds over 10 years ago, alternative data is now being sought for use in the wider asset management space, as well as the private equity and corporate verticals.

In parallel, there is an explosive increase in the supply of external datasets, as many corporates are looking to monetize their data and new technologies enable the emergence of new data vendors.

Eagle Alpha was one of the first companies to recognize the value of these new data sources and has been investing in educating and connecting alternative data vendors and buyers since 2012, in the process of building trusted relationships with both sides of this market.

As of December 2022, Eagle Alpha has profiled over 1,750 datasets and provides platform and data delivery solutions to data buyers and data vendors globally.

Eagle Alpha’s solutions mirror the user journey of our customers. There are three steps in the vendor user journey: discovery phase, productization, and go-to-market. There are three steps in the buyer user journey: data strategy, discovery and prioritization, delivery, and insights.

Our unique breadth of datasets, knowledge of the industry, and customer relationships have cemented Eagle Alpha as the global leader and strategic partner in the data space and the one true source of external data.

Eagle Alpha partners with industry leaders to continue to shape the industry.
1. J.P. Morgan, lead sponsor of our data conferences.
2. FISD, a member of this association creates standards for the industry.
3. Schulte Roth & Zabel, partner with this US law firm.
Contributors:

Niall Hurley  
CEO

Brendan Furlong  
Head of Data Sourcing & Advisory

Ronit Koren  
CMO

Thomas Combes  
CTO

Dallán Ryan  
Content & Advisory Lead, Data Strategy

Mikheil Shengelia  
Research Analyst

Aidan McDonnell  
Data Sourcing & Product Development

Ciaran Ryan  
COO