HFMTechnology examines how alternative data techniques are becoming more accessible to smaller managers

BY ANDREW MANNERS

Earlier this year Jiekun Huang, a softly-spoken and bespectacled assistant professor at the University of Illinois, uploaded a research paper to the website SSRN. Using a dataset of 5.9 million Amazon reviews between 2004 and 2015, he found consumer opinions contain key stock picking information. With abnormal scores predicting revenue and earning anomalies compared to an existing benchmark, a portfolio using this method can gain an extra 72-80 basis points a month, he wrote.

Huang also saw evidence that several hedge funds are already exploiting reviews correlated to the trading of some funds – especially those with shorter investment horizons.

Still a working paper, he could scarcely have imagined the buzz his research would attract.

In the days that followed, Huang was inundated with emails from hedge fund managers wanting to know more. *The Wall Street Journal* also wrote on his findings and earlier this month, he gave a talk to dozens quant analysts at Cubist Systematic Strategies, the systematic arm of Steve Cohen's Point72. Huang's findings highlight just one way that asset managers are leveraging Big Data.

While Wal-Mart founder, Sam Walton, would take to the skies in a two-seater jet to survey his firm's real estate investments, today firms can purchase satellite images to glean manufacturing trends, foot-falls or global trade patterns well before official figures are released.

Many of the industry's most successful quant funds, including Renaissance Technologies, DE Shaw, and Winton, have been harnessing Big Data to gain an edge for years. But as technology continues to advance at speed, such techniques are becoming accessible to many smaller managers too.

"What we are seeing is that with any company you talk to, the interest is there because they see the potential opportunity and also the cost if they are not able to keep up," says Balazs Klemm, founder and CEO of Almax Analytics, a data analytics vendor that uses AI to generate detailed reports on breaking news.

"Right now you have to have a Bloomberg terminal just so you don't lose out. I believe that this will soon be the case for Big Data," he adds.

The unstoppable rise of Big Data

Company filings, earning calls, and quarterly results have long been the core of sound investment analysis, but in recent years Big Data techniques have become another important weapon in a manager's armoury.

"In 2009 I was living in Silicon Valley and going to the Berkeley labs all the time. Seeing what was happening, I decided to leave [my job as co-founder of JP Morgan's Big Data team] to start a hedge fund," says Michael Beal, the CEO of $100m Data Management Capital, a fund that leverages novel data to trade US equities.

"People then were starting to ask different questions with new data... so even back in 2009 you could see something like this would come to fruition."

Powered by an online revolution and the internet of things, the growth in data has been unprecedented. According to estimates, 90% of the world's data was created in the last two years alone. Some of this data is structured in the form of market reports, official releases, or curated information, but nearly 90% of it is not posing challenges for asset managers.

"The information overload in capital markets makes it hard to make sense of things and process it in a timely manner," remarks Klemm. "Big Data has gone from not enough data to too much data."

But Hungarian-born Klemm, whose Almax Analytics is planning to launch later this year, says data harnessing techniques can be hugely beneficial as long as managers are able to filter out the irrelevant noise. For quant funds Big Data offers greater capability to help generate trading signals, create circuit breakers on bad news, and pull quotes rapidly on market-moving events.

For discretionary funds, meanwhile, it is slowly becoming a crucial element of researching global instruments and filtering out the noise of unstructured data to derive key indicators. Big Data is also being used by some funds to enhance market surveillance and improve risk management through forecasting volatility and liquidity. "Even if you're not finding alpha or improving your research there are benefits," declares Klemm.

Funds adopting Big Data

Despite this, though, few funds have fully tapped the potential of Big Data. "We estimate that there are approximately 50 funds worldwide today that have fully integrated altenative data into their investment process," says Emmett Kilduff, founder and CEO of Eagle Alpha, a firm that offers research, analytics, and data sets to asset managers keen to take advantage of Big Data.

> "
> Right now you have to have a Bloomberg terminal just so you don't lose out. I believe that this will soon be the case for Big Data
> Balazs Klemm, founder & CEO, Almax Analytics

"At least 95% of asset managers globally have not started integrating alternative data into their investment process," he says.

Kilduff, who grew up in Dublin and worked for investment banks in London, was inspired by a hedge fund manager to set up Eagle Alpha based on a simple premise: "That the amount of alternative data being created globally is growing exponentially and most of the buy-side, frankly - back then and today – are not taking advantage of it."

Frankfurt-based start-up manager Catana is one smaller firm looking to take advantage of alternative data. However, the bespoke €6m ($7m) systematic shop, which launched in June this year, uses Big Data to trade equity l/s and futures on the German stock exchanges. The brainchild of a number of researchers who began analysing social media eight years ago, they soon turned their attention to finance.

The fund's CEO and CRO, Holger Knauer, partnered with the researchers, who are based out of Bonn, when he realised the opportunity. "We don't know of any funds like ours that are solely based on Big Data," he says.

He adds: "Nowadays you can store and extract incredible amounts of data, but the problem is how to structure it, and many people don't have the expertise. We are just four people but we have the ability to collect all data that is published about the German stock exchanges. The system does this by crawling, understanding, getting a positive or negative read, then executing. The only discretionary thing we can do is to turn off our system, which we did before Brexit."

They were up 13.3% in the first half of 2016 prop trading on their system. The DAX, by contrast, was down -9.9%.

While the firm's Big Data Catana fund was down -0.8% in July, its prop trading results and novel strategy has attracted institutional attention.

"Given we have no established track record as a hedge fund, the interest from institutional investors is reasonably high," Knauer says.

Still, analysts believe most funds can do more to fully leverage Big Data. That is somewhat slightly perplexing given the barriers to using it have fallen so dramatically.

As Adam Broun, COO of Kensho, a data analytics firm that boasts the world's largest event data base and services top banks and hedge funds, says: "What makes the world different today than it was five or ten years ago is that the technology has moved to a point so that the way you can extract signal from the data is radically faster and cheaper and more flexible than it used to be because of new data and analytical technologies."

In particular, advances in distributed technology and natural language processing, or NLP, have proved to be a game-changer. Huang's Amazon equities research is a

case in point. He told *HFMTechnology* that all of his data crawling – all 5.9 million reviews on Amazon – was done by a few shrewd research assistant, over the course of a year, using NLP on the programme Python.

"Once you've set up the programme it should be relatively easy," Huang says.

And with more data, analytics providers and programming tools than ever before, it is estimated that Big Data financial analytics will be worth over $3bn by 2018, driven by growth in Big Data and cloud computing systems, both of which are worth close to $130bn and are growing at 23 and 30% per year respectively.

Yet it is not always easy to incorporate Big Data into your business model, and doubts persist around its costs and benefits (see p. 20 for this month's survey on Big Data).

Betting on Big Data

Ultimately, whether a fund should harness Big Data – and to what extent – will depend on their business model and operations. Depending on the resources of a fund, there are several options.

These include: licensing third party data sets that a team can mine; purchasing ready-to-use analytical tools (Eagle Alpha, Almax Analytics, Kensho etc.); or creating an internal data gathering system that can crawl the web. Firms can also use a mixture of these to save time and money.

Reluctant managers can begin by buying licensed data and deciding whether to establish internal teams later. This can be more cost-effective than gathering data internally. It will also be quicker to get start leveraging existing data, some of which may be structured, rather than crawling the web. So far, licensed data sets in hot demand include consumer transactions and web and app usage, but international data sets (outside of the US) and B2B data are gaining interest.

Meanwhile for managers wanting analytical tools from the outset, there is no shortage of vendors offering solutions.

As well as providing a thought leadership offering and a directory of alternative data vendors, Eagle Alpha sells analytical tools, including its Web Queries programme which allows mangers to submit bespoke requests for any publically available source on the web. The London-based company charges $125,000 per annum per team for all its data, research and analytical tools.

The thought leadership offering and directory are available as standalone packages for $30,000 and $40,000 respectively.

News sentiment and analysis provider RavenPack will set managers back around $75,000 to $200,000 for five users per year, depending on the product purchased. That is around the same as social media analysis firm MarketProphit, which charges between $3,000-$15,000 per month per user.

Kensho, meanwhile, was tight-lipped about a possible offering for smaller hedge funds but as development continues, do not be surprised if an option hits the market.

Finally, for those with the resources and knowhow, building an in-house capability is an enticing option. While few funds are able to invest in a dedicated data team, doing so can reap substantial rewards, particularly with funds able to consume data as they please. $13bn New York hedge fund giant Tudor Capital is one firm which has been investing heavily in this area in recent years, hiring a team to look at how this technology can boost investment and operational alpha.

"You've got to find your own edge; if you're going off and buying analytical programmes or data and trading off the back off it, then someone else is likely to be doing it," says Andrew Flatt, CTO of $925m Omni Partners.

Funds building an in-house capability must also attract smart people with deep IT knowledge, something banks have struggled with following the financial crisis.

"You have to attract people that don't like finance. We've had to adapt our culture to bring in those people," says Beal, whose firm hires PhDs from different disciplines and aims to collaborate with other companies. "Any company that is commercialising data via an API, we want to speak to."

> “
> You've got to find your own edge; if you're going off and buying analytical programmes or data and trading off the back off it, then someone else is likely to be doing it
> Andrew Flatt, CTO, Omni Partners

Big name hedge funds with significant resources are steadily moving into this area. In March 2015, Steve Cohen's $12bn family office Point72 announced the launch of Aperio, a newly-minted Big Data division with 30 data scientists. Months later in July, it hired Bridgewater's former head of research, Mathew Granade, to head up its data and analytics team.

And earlier this year, $30bn systematic giant Winton announced it was building out a data centre in San Francisco it hopes will host 40 data scientists.

The firm is aiming to attract top local talent from Berkeley, (where founder David Harding has strong ties), and Stanford-both are considered amongst the best universities in the world for mathematics.

"All of the best funds are incorporating Big Data into their strategies," says Beal, with a number now attempting to use Big Data and AI to build machines that can predict market trends and trade on them autonomously, rather than producing novel patterns for managers.

From a €6m Frankfurt start-up to $150bn giant Bridgewater, funds around the world are waking up to the growing potential of Big Data.

"Disruption is coming to the capital markets," Beal predicts. "Those who can adapt to this will do extremely well, while those who don't will have some issues in the next few years."