Hedge funds

Scraping data: the next frontier in battle for investment returns

Use of consumer emails and receipts is growing but brings potential problems

The flood of data scraped from emailed receipts and other online activity can be used to help funds forecast companies' performances © FT montage

Robin Wigglesworth DECEMBER 11 2016

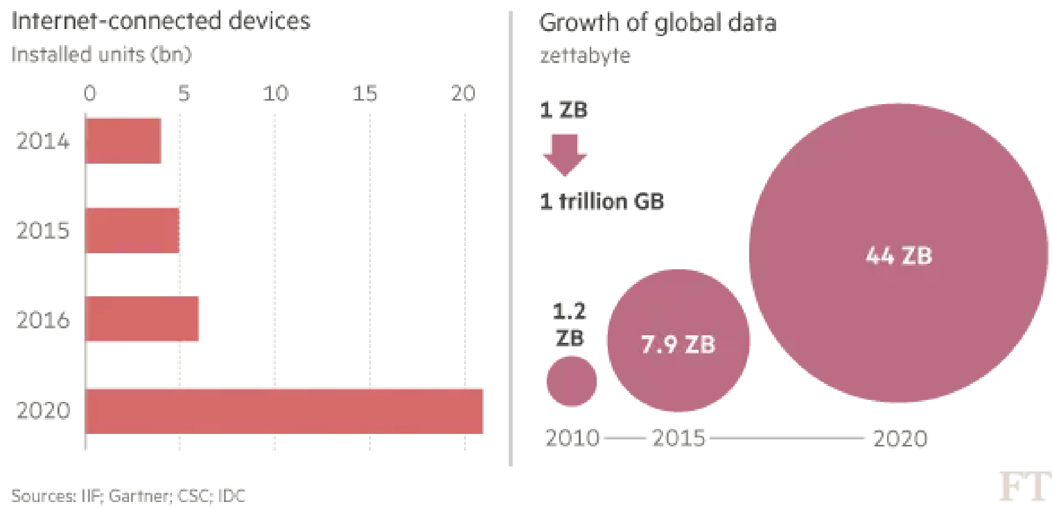
A money manager might once have dispatched a junior grunt to the local Sears department store to gauge how sales were doing. This year a hedge fund could be peering into your emails for hints on how Amazon, the world's largest online retailer, will do in the critical Christmas quarter.

One of the latest technological weapons deployed by the most sophisticated asset managers desperate for an investment edge involves doing just that, albeit in a roundabout way.

There are technology firms that offer companies the ability to scan incoming emails. However, software "cookies" can also look for emails sent from the likes of Amazon, Walmart or Best Buy, check for sales receipts and then log that data in a database. That data then lands in the lap of hedge funds via a third-party vendor.

Scouring emails for such nuggets of information is just the tip of the "alternative data" iceberg. Almost everything we do leaves a digital fingerprint, which can be scraped, aggregated and sold to investment firms looking for tradable signals, or to use the jargon — market-beating "alpha".

“The buyside is struggling more and more to find alpha, but data is alpha, and more and more of it is becoming available ever day,” says Tammer Kamel, the co-founder and chief executive of Quandl, a US alternative data provider. “The demand is there, the supply is now there as well, and we want to sit in the middle of this.”



Supply is certainly swelling. The world’s annual data generation will more than double annually for the foreseeable future, and reach 44 zettabytes (trillions of gigabytes) by 2020, according to IDC. To put this in context, if all this information was put in iPad Air tablets, the stack would reach from earth to the moon more than six times over.

Luckily, computing power has kept pace. Today’s high-end Nvidia video cards are so powerful they would have been classified as the world’s most powerful supercomputer before 2002, Goldman Sachs analysts recently noted.

Coupled with huge strides in artificial intelligence — particularly in areas like “deep learning” — this is allowing investment groups to mine enormous data sets to unearth insights they hope will keep their returns ahead of rivals.

Take scanned email receipts. Earlier this autumn, GoPro, the action camera maker loved by daredevils, reported a terrible quarter of sinking sales, sending its shares tumbling. Yet by using a database of electronic receipts extracted from over 3m email inboxes, Quandl says it could see ahead of time that revenues would disappoint badly.

The potential of alternative data has long been known to hedge funds, but traditional asset managers, looking for ways to improve performance, are now jumping in. For example, UK asset manager Schroders has set up a “Data Insights” team to “enhance and complement the existing skills of our fund managers and analysts”.

The analytics company Quandl said it was able to anticipate GoPro's recent poor results using its receipts data © Reuters

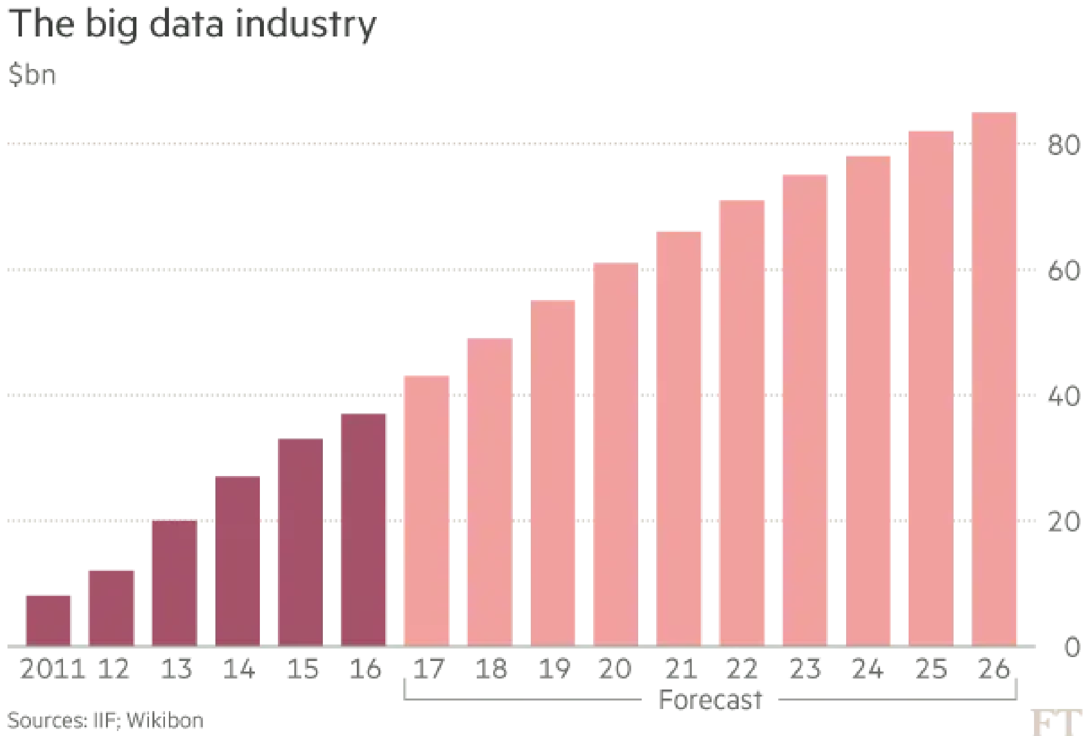
"Global trends in digitalisation, social media, open data and technology are all creating vast streams of alternative data that are often highly unstructured and extremely obscure. However, they contain valuable and often unique insights," the investment group's latest annual report said.

The investment industry's push into alternative data raises some potentially awkward questions. First, some money managers are concerned the personal data are not always scrubbed completely from the data sets they are sold. The majority of vendors do a good job, but standards are far from uniform, they say.

"We talk about it all the time," said Wesley Chan, director of stock selection research at Acadian Asset Management, a $70bn Boston-based firm. "On one hand, we're excited about the potential, on the other we're racing to delete apps on our phone. The data vendors claim to anonymise the data through aggregation, but it doesn't always work perfectly."

Mr Kamel stresses that Quandl, one of the firms with a better reputation, is obsessed about scrubbing off personal information and says there is a "healthy paranoia" among the company's clients about the issue.

"No one wants to be on the wrong side of this," he says. "We make it our mission to mitigate 100 per cent of those risks. One fiasco in this space and the consequences for a fund would be dire."



Scraping, selling and buying data are perfectly legal in most countries, as long as safeguards are taken. Even so, Mr Chan warns mishaps are probable. “Six months ago I would have said we just have to live with it, nothing is private any more anyway. But at some point there will have to be some sort of bill of rights to control it,” he says. “This will continue until there’s a scandal, which is likely to happen.”

If fund managers begin to lean more on alternative data to help make investment decisions, some believe it ultimately has the potential to dislodge quarterly corporate earnings reports and monthly economic data from the centre of investors’ radars.

So far big data can only offer clues — some strong, some weak, some simply wrong — on what these market-moving releases will show. But in time, it is far from inconceivable that our digital fingerprints will provide an accurate, instantaneous gauge that will render traditional investment information obsolete.

“As this perfect storm of technological sophistication comes together, we have never been closer to charting out a real-time map of consumer behaviour in both the real and digital world,” Tabb Group, a research firm, noted in a report on the industry.