

The Globe & Mail

[The data game: How information on everything from flight patterns to parking lots can reveal valuable clues about where the market is heading](#)

By Tim Shufelt
April 19, 2019



Last summer, a corporate jet owned by Encana Corp. embarked on a strange pattern of flights.
THE GLOBE AND MAIL

Last summer, a corporate jet owned by Encana Corp. embarked on a strange pattern of flights.

Over the course of a couple of months, the Calgary company's aircraft was tracked landing close to oil fields in Oklahoma, Utah and Montana. Encana had operations in none of those places. Houston-based Newfield Exploration Co., however, owned assets in all three locations. And it turns out that Newfield's jet was also on the move around the same time. Its executives seemed to take a particular interest in flying to Denver, which happens to be home to Encana's chief executive officer, Doug Suttles.

The mutual visits hinted that something was in the works. And in early November, Encana struck the largest deal in its history when it agreed to buy Newfield for US\$5.5-billion. An aggressive expansion into U.S. shale represented a dramatic change of course for the Canadian

driller, and the deal caught the market by surprise. When trading started the next morning, Encana's stock plummeted by 17 per cent. A select group of investors, however, may not have been quite so shocked. They had access to the movements of Encana's jet, giving them strong clues about what the company's executives were up to in the months before the acquisition.

Quandl Inc., a Toronto-based startup, allows investors to peek in on the travel habits of companies such as Encana. By stitching together data from aircraft registries, corporate filings and flight communications, Quandl can track the movements of thousands of corporate jets around the world, giving investors a new advance signal of potential market-moving corporate deals.

Known as "alternative data," this kind of insight into publicly traded companies is proliferating, as investors and fund managers look to data science for an edge that will help them beat the market. Jet activity can foreshadow corporate deals; aggregated credit-card tallies can reveal consumer trends; satellite imaging can track oil inventories; information "scraped" from job sites can indicate who's hiring; data collected by auto insurers can give clues on future car-sales figures. The vast pools of information being generated by the digital economy hold the power to better predict what companies will do and how their stocks will perform.

The computer sophistication and machine learning needed to make sense of all that information, meanwhile, is quickly evolving. Until recently the purview of quantitative hedge funds, alternative data methods are spreading to mainstream investing. In December, Quandl was acquired by Nasdaq Inc., which runs a data products business in addition to operating stock exchanges. The deal amounted to an "inflection point for the industry," wrote Richard Johnson, a vice-president in Greenwich Associates' market structure and technology group. Competition among data providers is heating up. In February, Bloomberg launched a site featuring 20 alternative data sets. "The race to take alternative data mainstream has now begun in earnest," Mr. Johnson said.

It's a race that Canadian investors have been reluctant to join. Despite its Toronto address, Quandl has zero Canadian names on the list of major clients subscribing to its 50-odd data products, which range from estimates of Tesla sales to industrial-auction results. Outside of the big pension funds, in fact, it appears few Canadian investors are dabbling in alternative data.

There are reasons to be cautious. There is a growing discomfort over the capacity for Big Data analytics to observe the intimate details of people's lives. Meanwhile, the legalities over data collection and distribution can be murky, raising concerns over who owns particular information and who has the right to sell it. There are tough questions regulators are just starting to grapple with, including whether sophisticated investors gain an unfair advantage when they have access to data that is effectively unavailable to the masses.

And yet, wary Canadian investors run the risk of being left behind if they wait too long. Alternative data will soon be essential to generating competitive returns, says Tammer Kamel,

Quandl's CEO. "It will become unacceptable to be basing your investment decisions on what happened a few months ago."



Co-founder and chief data officer Abraham Thomas, left, and CEO Tammer Kamel, of Quandl, during an office meeting in Toronto on Feb. 13, 2019.

TIJANA MARTIN/THE GLOBE AND MAIL

For decades, standard financial data has been the lifeblood of fundamental investing. Investors glean what they can from whatever public companies are required to disclose through regulatory filings and quarterly financial statements. A handful of data providers, including Bloomberg, Refinitiv and FactSet, have come to dominate the distribution of that information. (Refinitiv is partially owned by Thomson Reuters Corp., which is controlled by Woodbridge Co. Ltd., owner of The Globe and Mail.) In 2018 alone, investors spent in excess of US\$30-billion globally for access to market data and analysis, according to an estimate by Burton-Taylor International Consulting.

Alternative data means anything considered to be outside the realm of traditional financial information, but that can yield valuable market or company insight. And it's by no means a newly invented category. Investors have long hunted for tradable information outside the bounds of financial reporting. It used to be said that the thickness of U.S. Federal Reserve chair Alan Greenspan's briefcase could portend monetary policy announcements. (A big haul meant he was carrying the documentation to support a rate cut, or so the theory went.) Hedge fund managers have also been known to directly observe retail foot traffic or cross-border shipping

or executives appearing in certain airports – anything to get a read on what, or how well, a company is doing at that very moment.

What has changed is the sheer volume of data now being produced, everywhere. The internet has more than 1.5 billion live sites. Facebook users create about 3.3 million posts a minute. The Internet of Things is connecting everything from cars to household appliances, and smartphones are constantly tracking their users' locations. By next year, roughly 1.7 megabytes of data will be generated each second for every person in the world.

Most commercial information is simply “exhaust” – a byproduct of a company's main business, Mr. Kamel says. But there is an active market for those companies to turn their data into revenue. Many telecommunications companies give third parties access to user location data for a fee; financial intermediaries will compile credit-card transaction data; and policy information from auto insurers can reveal which models of cars are selling best. “You can find out almost anything you want to know about a stock or a commodity or a consumer, if you connect to the right database,” Mr. Kamel says. “Somebody's taking that measurement.”

That's where alternative data providers come in, typically licensing that information and turning it into data sets marketed to big hedge funds and asset managers. There are currently more than 400 providers like Quandl, up from around 100 a decade ago, according to alternativedata.org. “Web scraping,” or data extracted from websites, is the largest subcategory. Using data scraped from Best Buy Co. Inc.'s website, for example, New York-based startup Thinknum showed robust sales in Amazon products, such as its Alexa-powered smart speakers, starting around Black Friday last year – nearly one month before Amazon.com Inc. announced record holiday sales for its devices. Research firm Opimas estimates that hedge funds and asset managers scraping sites for investment purposes accounted for 5 per cent of all web traffic last year.

So-called sentiment data scraped from social media, financial news and online forums are among the more established alt data products. Toronto-based Buzz Indexes built a model that scours sites like StockTwits and Twitter for insight into how investors feel about individual stocks. A natural language processing algorithm looks for signs of investor positivity toward U.S. large-cap stocks and calculates a sentiment score for each name. The 75 stocks with the highest scores are included in the Buzz NextGen AI US Sentiment Leaders Index, which, back-tested to the start of 2013, has returned an average of 17 per cent annually, compared to 11 per cent for the S&P 500 index. Not too long ago, the idea that there might be wisdom in the online conversations of investors was met with cynicism, says Buzz Indexes founder Jamie Wise. “Today, there's probably not a CIO at any major asset manager on the continent that isn't thinking about an alt data strategy.”

Many have progressed well beyond the thinking stage. BlackRock uses an active quantitative approach in its Advantage funds, which search for investment signals from a range of data sources, including weather patterns, travel-site bookings and employee reviews from sites like Glassdoor.com. “Combining millions of responses can indicate a company's state of health, as

those with happy employees tend to outperform their competitors,” BlackRock said in a recent brochure for its Advantage funds. Meanwhile, Franklin Templeton Investments recently signed a deal with platform company Elsen, giving traditional portfolio managers easier access to big sets of data.

For large investors and asset managers, getting access to market and company intelligence that gets as close as possible to real-time data is worth paying good money for. Quandl’s datasets range in price from US\$25,000 to US\$250,000 a year. Other products on the market, like specialized satellite intelligence, can cost upwards of US\$1-million a year.

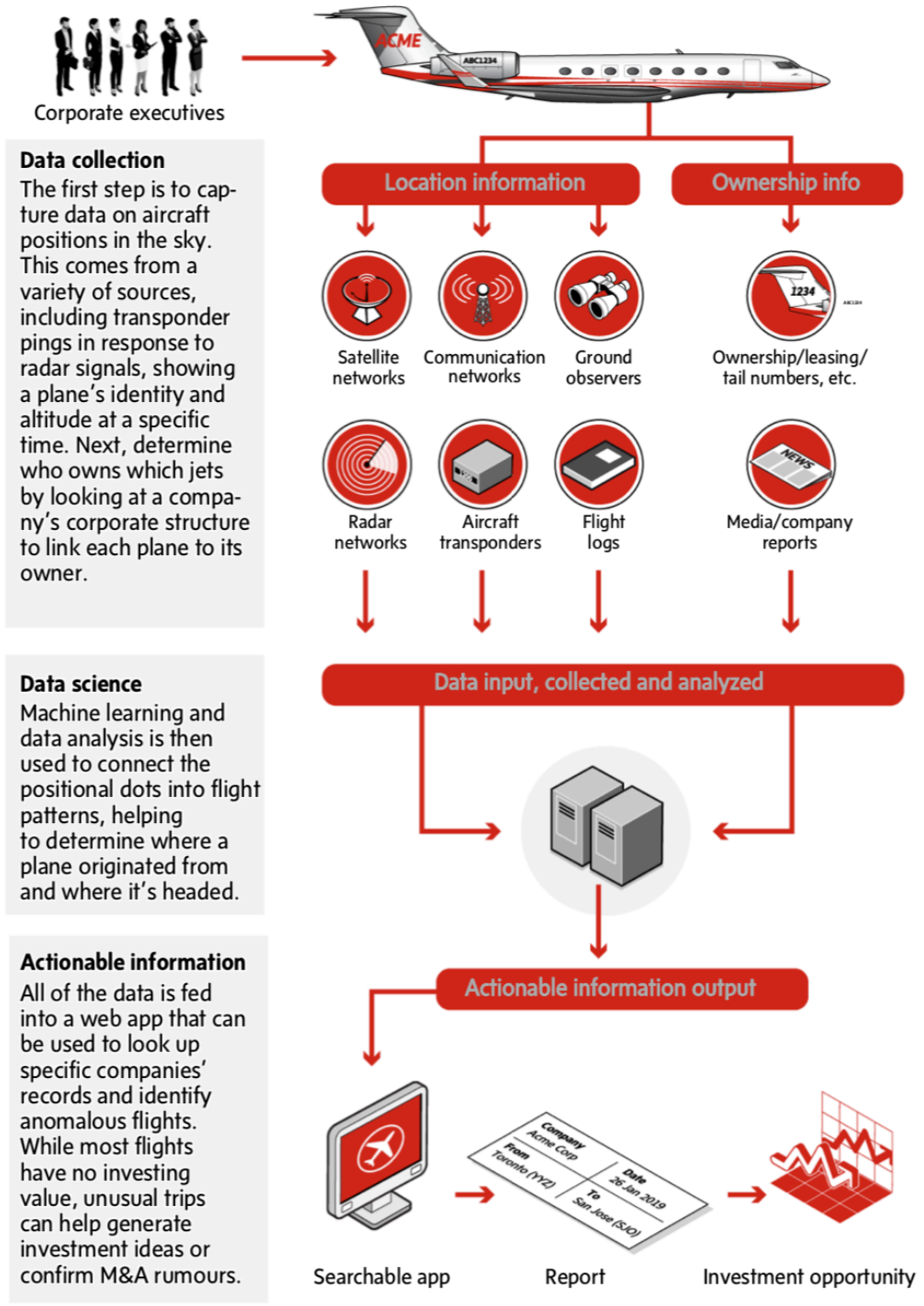
A pair of professors from the University of California at Berkeley recently demonstrated how satellite images’ predictive power can justify such an extravagant price tag. They looked at daily images of the parking lots of major U.S. retailers, including Walmart and Target, over a six-year period to identify whether counting car traffic could help predict earnings and stock movements. A trading strategy built on buying shares in retailers with abnormally high parking lot traffic, and shorting those with low traffic, would have paid off handsomely once earnings were announced, the analysis found. Compared to a buy-and-hold approach, that satellite-informed portfolio generated average excess returns of 4.6 per cent.

Silicon Valley-based Orbital Insight is one of the leaders in using satellite technology to spot tradable economic or company data in real time, mostly in the consumer and energy sectors. Last September, RBC Capital Markets announced a partnership with Orbital that provides the investment bank with geospatial data to include in its equity research. An RBC report from January said Orbital’s images of storage tanks pointed to declining global crude oil inventories from their December peak – one sign the market could be tightening and prices headed higher.

“Everybody’s trying to get into the alternative data space,” says Fardeen Khan, head of strategic initiatives at RBC Capital Markets. But he adds that it’s not a standalone investment approach. The idea behind RBC’s arrangement with Orbital, as well as the bank’s other data science endeavours, is to complement the fundamental and technical process. “When you look at alternative data as a standalone, the insights are not sufficient to say you should go fully long on this company or go short on a specific name,” says Mr. Khan.

That sentiment is echoed by Ron Mock, CEO of Ontario Teachers’ Pension Plan, which uses some data-driven trading strategies and is “leveraging the deep insights it’s capable of bringing,” he said during a discussion at the World Economic Forum in Davos, Switzerland, in January. “We have to be very, very mindful, that we can’t push it so far that we turn our brains off.”

Quandl's corporate aviation data product provides investors with the ability to track the comings and goings of corporate executives. Where they fly can often give hints about what companies are up to – who they're meeting with, where they're focused on, and what big deals they might have in the works.



Nearly a decade ago, hedge funds were the only ones most willing to take a chance such an exotic, untested idea as using alternative data, and they have been the main driver behind the growth of that industry. Global spending on alternative datasets is currently about US\$3-billion per year, according to JP Morgan, a small fraction of the size of the conventional data business.

For the industry to assume a larger profile, it will need to extend its appeal to more traditional asset managers. Now is a good time to do just that, Greenwich's Mr. Johnson says. "A lot of active managers are struggling to beat passive benchmarks, and they're looking for a new edge," he says. The passive investing craze has made life difficult for traditional active managers. Franklin Templeton, for example, saw its global assets under management decline by 14 per cent last year.

The average active fund manager, however, has very different data needs than a giant U.S. quant fund. Without the infrastructure to analyze raw data, most fundamental investors require data that have been ingested, formatted and packaged, or fed into platforms they can incorporate into their own investment processes. Quandl's corporate-jet-tracking app is one example of this.

The idea for an aviation-based investment tool came out of a hedge-fund trade from early 2017. A trio of New York-based funds figured out how to track Johnson & Johnson's Gulfstream jet on the internet and found it sitting on the tarmac at a Swiss airport for more than a week, just a few kilometres away from the headquarters of pharmaceutical company Actelion Ltd. Convinced a major tie-up was being negotiated, the hedge funds loaded up on Actelion shares, which soared when J&J announced a US\$30-billion deal to acquire the Swiss company a few days later. When Abraham Thomas, Quandl's chief data officer, read about that payday, he thought: "What if we could automate that process?"

By combining flight location data and ownership information from several different sources, Quandl can now track the flight activity for a majority of the companies in the Russell 1000 Index. Most of those companies, however, would prefer to keep that information to themselves. They'll often try to conceal their own aircraft through subsidiaries and holding companies, or complex leasing arrangements. By poring over aircraft registrations, operator licences, public filings and corporate parent-subsidiary relationships, Quandl has built a database of 29,000 jets and counting.

Most subscribers are using the product as one part of an M&A investing strategy, to help shed light on rumours or suspected deals, Mr. Thomas says. Others use it for protecting their short positions. Investors betting against a stock are vulnerable to that company being acquired, since such an announcement typically results in a big jump in share price – and big losses for short sellers. "On other occasions, activist hedge funds want to find out if the CEO is gallivanting around the world on the company dime," Mr. Thomas says.

Corporate jet data is one of dozens of data sets that Quandl says puts it in the alternative-data lead, Mr. Kamel says. Already, the company leads the industry in brand recognition, according

to a recent Greenwich Associates study. And being acquired by Nasdaq represents a huge boost to the company's profile and credibility. "When you hand someone a card that says Nasdaq – part of the fundamental structure of capital markets – that helps a lot," Mr. Kamel says. "Now we're standing on the shoulders of a giant."

So far, Canadian hedge funds have taken a pass on alternative data – almost all of them, in fact, according to Claire Van Wyk-Allan, head of the Canadian chapter of the Alternative Investment Management Association. Most Canadian players are just not big enough to justify the cost. The hedge fund industry here pales in comparison to behemoths on the other side of the border. Bridgewater Associates, for example, manages about US\$160-billion, while only a handful of Canadian hedge funds surpass even the US\$1-billion mark.

"We are not able to spend \$100,000 every month on all kinds of data. We aren't Bridgewater," says Ernest Chan, who runs QTS Capital Management in Niagara-on-the-Lake, Ontario, and manages a small hedge fund. His own experience with alternative data suggests it generates one to two percentage points of "alpha," or excess returns. For Bridgewater, that would amount to a boost to annual returns of US\$1.6-billion to US\$3.2-billion. "But if you are \$100-million fund, alternative data is not must necessarily a must-have," Mr. Chan says.

The major Canadian pension funds, on the other hand, are certainly big enough to use algorithmic trading strategies and advanced data analytics. "If you go to a quantitative investment conference, it is dominated by pension plans," Mr. Chan says. Canada Pension Plan Investment Board, Ontario Teachers' Pension Plan and Alberta Investment Management Corp. all declined to comment on how they're using alternative data in their investment decisions.

Canada's big asset managers, meanwhile, appear to be on the sidelines when it comes to alternative data. While Quandl boasts of having 14 of the world's 15 largest asset managers as customers, the company has yet to land a big Canadian name. "It's a little frustrating that it's easier for me to sell in New York than in my own backyard," Mr. Kamel says. Though Toronto has emerged as a global fintech hub, Bay Street asset managers seem reluctant to evolve. "Canadians just might be used to doing things the regular old way," Ms. Van Wyk-Allan says.

Part of the hold-up might be in how alternative data is typically marketed. "Looking at this data as a source of alpha is like entering into a nuclear arms race, where you're constantly in search of the next data set," says Ashby Monk, executive director of Stanford University's Global Projects Center, which studies, among other things, how technology can improve long-term investing. "If you're a patient investor, that's probably not the best use case for alternative data." Instead, he suggests it can be used to better understand risk in a portfolio, to assist in due diligence and to make better capital allocation decisions.

Alternative data is by no means risk-free, however. It's an unregulated space that lacks legal clarity. While insider trading generally has a fairly narrow legal definition, some alternative data strategies are starting to look awfully similar. The common regulatory test for insider trading asks whether a piece of information is material and non-public. Alternative data certainly has

the power to be material to a company's stock, by providing timely indicators of a company's health. And if distributed only to a limited number of investors, or even exclusively to a single hedge fund, it can be difficult to argue that data is public.

"The line between public and material non-public information is key here," says Kirsten Thompson, a partner at Dentons and national lead of the law firm's transformative technologies and data strategy group. "Securities regulation requires you to know which side of that line you are on, and alternative data makes that difficult." Some large hedge funds are known to avoid buying "exclusive" datasets for fear of legal risk.

There is growing alarm, meanwhile, around the sharing of personal data by companies that compile it. Last year, it was discovered that location information on U.S. cellphone users sold by telecom providers was ending up in the hands of bounty hunters. Canadian telcos also sell location data to third parties, which they say is only done with users' explicit consent.

Canadian privacy legislation regulates the personal information that corporations collect, limiting the ways it can be shared. And in any case, investors are not interested in obtaining anyone's personally identifiable information, Quandl's Mr. Thomas says. "We tell our vendors, if you have personal records, don't even send us the data. We don't even want it to touch our servers."

But it can be difficult to truly anonymize certain information. "Our machine learning and our algorithms are now getting so sophisticated that you may have data that you think is not identifiable, but the algorithm can, in fact, identify somebody," Ms. Thompson says. Data science can also be used to approximate protected data. Access to Canadian credit scores is limited, for example, but a fairly accurate estimate can be computed from other sources, including social media.

Investors considering accessing alternative data need to conduct their due diligence, Ms. Thompson says. "There's a constellation of questions you should be asking, including the sources of data, appropriate consents, the genesis of the information." Data harvesting methods like web scraping can violate the terms and conditions set out on a website's fine print, in which case, data vendors would be prevented from selling that information.

Quandl says it is offered about 100 data sets each month by companies and data hunters looking to cash in on the data they have collected. In addition to evaluating the data quality and its potential value to investors, the company says it takes pains to trace each potential data set back to its original source. "Our customers understand we've vetted the data for personal information, insider information and ownership issues," Mr. Thomas says.

Outside of Canada, adoption of alternative data is growing fast, despite the regulatory limbo. Canadian investors won't have the luxury of taking a cautious approach much longer, Mr. Thomas says. "As it spreads wider and wider, if I don't have this data, I'm at a disadvantage. It becomes table stakes."