



Analysts Beware! A Machine Has Its Eye on Your Job

23 OCT 2013 - S.L. MINTZ

BIG DATA IS SUDDENLY A BIG DEAL. In late May, at Michael Milken's annual gathering of financial notables in Los Angeles, one center-stage topic was the impact of increasingly powerful computers and software and their ability to extract meaning from giant data sets. The conference featured one panel, headlined by President Obama's 2012 campaign manager, Jim Messina, that focused exclusively on big data; Messina discussed how vast amounts of information helped secure victory. But it was a panel on trading that defined the challenge to flesh-and-blood financial and investment analysts posed by the seemingly inexorable march of algorithms, robots and big data.

Louis Salkind, longtime managing director of hedge fund firm D.E. Shaw & Co., framed the challenge of big data most clearly: Salkind, who has a Ph.D. in computer science and robotics from New York University, described a mounting confrontation between analytical machines and securities analysts. Near the end of the hourlong panel, he described a world of increasing automation in which vendors create products "to shred apart Twitter and Facebook" and aggregate trading signals. "Imagine what happens when they start using big-data techniques to look at fundamental data," he said, recounting a story about a broker who used satellite imagery of Wal-Mart Stores parking lots to forecast quarterly earnings. "When people start integrating these forms of data, it's just going to be a different world out there."

Salkind and D.E. Shaw have long been innovators in the use of computers for trading and investment. But some believe that the struggle between automation and human practitioners of finance, which has swept through exchanges, trading floors and even

regulation and compliance, has reached a new front: securities analysis.

Computers and software clearly have advantages in tracking complex market patterns or monitoring and analyzing data points in news reports, social media and other digital sources. Computers famously never have to use the bathroom or ask for a raise, though they do break down. They are growing increasingly fast and more powerful, and they have access to far more data. Apostles of big data predict a rout of rank-and-file analysts by computers that can interpret the market with superior results. They even believe that big data will allow machines to discern the future — for an election, a stock price or a corporation — from the noise of the moment. Big data will not only reshape trading, they say, but long-term investment practices.

Others are skeptical. The future is always unpredictable, no matter how much data is aggregated and analyzed, as long as people participate in markets and economies. So far, the practical performance results have been thin: We suffer flash crashes, unhappy financial shocks and bubbles hidden in plain sight. Human analysts cannot process the flood of data that machines can, but they possess something algorithms lack: finely grained, if fallible, judgment. Human analysts can weigh murky values that may not be reducible to quantification; balance long-term and short-term perspectives; profit from intuitions about companies and their futures; forecast the evolution of technologies, brands or fads; and cope with ambiguities and complexities.

This clash of man and machine is just the most recent chapter in a centuries-old struggle that heated up when mechanical looms replaced home-based spinning wheels. Although the outcome is not clear, what is obvious is that the world of securities analysis will change under the impact of these powerful new tools. The technology may well transform the already precarious economics of securities analysis and further cull the ranks of analysts, dividing them into those who can effectively use the new techniques and those who cannot. Big data is probably here to stay. The larger question is, how do we live with it?

At the heart of this trend is the algorithm, a series of steps or instructions that tells computers how to search for and interpret data. It's a simple but powerful concept when

allied with a computer. Consumers encounter algorithms every day. In addition to routine tasks, from spell-checking to GPS route guidance to online shopping, algorithms help fly passenger jets and perform medical diagnoses, even surgeries. Soon they may drive cars.

Algorithms are also ubiquitous in finance, playing a role in everything from high frequency trading to complex valuation calculations to economic forecasting. They feed off information that washes over global finance daily — data now measured in petabytes, or billions of megabytes. No army of humans could outprocess these algorithms in weeks, much less in the fractions of a second they need to churn through data.

Algorithms excel at performing rapid, nearly limitless computations. But they require data as raw material. That data is increasingly available in large quantities, much of it a world apart from the traditional grist of financial analysis: prices, valuations, ratios. Increasingly powerful computer systems squeeze market patterns from news items, financial statements, blogs and other digital texts where insights may lurk — a developing field known as news analytics. Massive digital memories keep tabs in real time on thousands of companies, along with their competitors, customers, vendors and investors.

News analytics weigh factors from arrays of financial ratios to CEO transitions to local work stoppages in rural China for individual stocks, sectors, peer groups or the market. Programs scan horizons and flag events with market implications. Some algorithms even pump out news updates for media consumption.

Fresh data can yield trends and new peer groups tied together by market sentiment, supply-chain relationships or news events. Better yet for investors, novel peer groups often exhibit unique trading patterns.

“Will this source of information change the investment industry? Undoubtedly, the potential is there,” says Wesley Chan, a finance Ph.D.; Goldman, Sachs & Co. veteran; and director of stock selection research at Boston’s Acadian Asset Management, a quantitatively based investment manager that uses news analytics. “Accounting and market data became very important to the investment industry over decades. There’s no

reason to think news analytics won't have the same impact in an even shorter time frame.”

What does this mean for securities analysts? Some advocates of big data believe computing power and predictive algorithms will sweep away traditional analysts and, by extension, a traditional approach to investing. The outcome depends on how effective some of these algorithm-based predictive systems prove to be and whether traditional investment research and active portfolio management can make an effective case for survival.

However, advocates of the continuing centrality of human judgment argue that the binary decisions at the heart of algorithms won't necessarily make the world more predictable and that the merits of long-term investing will persist. After all, every algorithm has a set of instructions devised by mere people; like economic models, algorithms simplify the world down to manageable inputs and outputs. Human experts command a wider variety of knowledge than algorithms and apply experience in ways that algorithms cannot, says James Owen Weatherall, an assistant professor of logic and the philosophy of science at the University of California, Irvine, and the author of *The Physics of Wall Street: A Brief History of Predicting the Unpredictable*.

Medical research furnishes a classic example. A 1972 study asked oncologists to predict, on the basis of biopsies, the survival time of 193 Hodgkin's lymphoma patients. The correlation between expert predictions and actual survival time was zero. But the coding of the biopsies generated by those physicians and run through a multiple regression model accurately predicted the survival time. To Weatherall this illustrates a key difference between human and automated roles. Researchers are better at identifying the variables that computers need to function. Computers can more effectively synthesize information in a systematic way to make predictions.

In finance, algorithms have played a growing role since the 1980s, principally in trading. But as D.E. Shaw's Salkind suggested, companies with proprietary algorithms have turned their ambitions to buy-and-hold investing and to fundamentals. Vendors such as Bloomberg, Dow Jones & Co. and Thomson Reuters have joined an ever-expanding

collection of fledgling firms with names like [Alexandria Investment Research and Technology](#), [AlphaGenius Technologies](#), [Dataminr](#), [Digital Trowel](#), [Lucena Research](#), [MarketPsych](#), [Narrative Science](#), [Quantopian](#), [RavenPack](#) and [Recorded Future](#).

These companies sell various twists on news analytics. The Big Three firms aggregate suites of interactive products. Eikon, marketed by Thomson Reuters, deploys news analytics in a user-friendly window on market trends, securities prices and even the exact location of oceangoing freighters on a digital map. In isolation a freighter's coordinates may not alter the outlook for a company, but in conjunction with news, weather conditions and market demand, a severe storm might affect market value.

Meanwhile, start-up AlphaGenius mines Twitter traffic for market signals. Recorded Future sweeps major news publications, trade publications, government websites and financial databases for explicit and implicit signs of future events. RavenPack offers data products and advanced visualization tools that identify countries or companies for which sentiment or media attention is high or low. In search of market patterns, Alexandria applies computer technology developed in human genome research. Lucena Research, founded by a former F-15 pilot with a Georgia Institute of Technology Ph.D. in robotics, couples artificial intelligence with investment strategy. Quantopian equips scientists with the means to develop and backtest their own financial algorithms.

TODAY'S ANALYTIC TECHNIQUES go far beyond the simple word counts that hobbled earlier forays into news analysis and attempts to track market sentiment. Those efforts labeled words positive or negative, anchored in the simplistic assumption that negative words always mean bad news and positive words always mean good. Context, long thought a strength of human judgment, has become paramount.

Windows for leveraging news analytics and market sentiment do not just flash by like an equity trade in a dark pool; they can, practitioners assert, last days, weeks, months or longer, thus becoming relevant to investment analysis. "I've seen clients trading on our information on investment horizons for up to three years," says Peter Hafez, director of quantitative research at RavenPack and a former portfolio manager. His customers seek quantitative frameworks for news that does not require interpretation by analysts. "As you

get stronger news analytics and more information, you can bypass analysts and find value in the data itself,” he says.

In February the *Harvard Business Review* published a case study that lent validation to sentiment detection at Recorded Future, an analytics start-up based in Cambridge, Massachusetts. The study reported on a Recorded Future strategy that categorized 500 stocks according to sentiment, then bought the top 10 percent and shorted the bottom 10 percent. “RF’s own analyses suggested that if investors had followed its predictions and investing recommendations about equities in the S&P 500 over the past year, they would have substantially outperformed the market,” according to the case study.

Academic research over several decades has tended to corroborate strategies that exploit news. “Stocks of low capitalization, younger, unprofitable, high volatility, non-dividend paying growth companies, or stocks of firms in financial distress, are likely to be disproportionately sensitive to broad waves of investor sentiment,” researchers Malcolm Baker of Harvard University and Jeffrey Wurgler of New York University reported in a 2007 paper, “Investor Sentiment in the Stock Market.” Other studies concur that sentiment signals can generate advance warning of market changes, although opinions vary on the duration and magnitude of impact.

To test the value and life span of news analytics, a series of recent research papers by Deutsche Bank focuses on state-of-the-art sentiment analysis. Quantitative strategist Rochester Cahan and his colleagues explore using news — known in the jargon as unstructured data — in stock selection. Their conclusion: The real value in news and Internet data lies beyond simple long positive–short negative sentiment strategies.

Sentiment in absolute terms has less meaning than sentiment relative to market expectations, the Deutsche team concludes. Successful financial models extract alpha from news by capturing complex interactions between sentiment and market data variables like price and volume. “If a company has lots of good sentiment, people writing good things on blogs, tweeting good things, there is an automatic assumption that that is a positive story and you should buy,” Cahan says. “Markets don’t work like that. What matters is expectation.”

Sentiment furnishes nuanced proxies for expectation. Value, bias and context all color sentiment, says E. Paul Rowady Jr., a senior analyst at TABB Group, which monitors news analytics. A layoff might be good or bad for sentiment, depending on its context. Some are buy signals; others, sell signals. Bad news for a Ford Motor Co. supplier might suggest good news for General Motors Co., or not. Strong earnings by a market leader might force rivals to scramble to maintain market share or, conversely, surf a rising tide. Programmed properly, news analytics algorithms can recognize implications and sentiment in context.

Eventually, Rowady foresees universally accessible data and computational firepower strong enough to absorb everything that is happening in real time and then express market sentiment. "That's essentially where we are headed," he says.

CAN NEWS ANALYTICS PRODUCTS predict the market future any better than traditional analysts? That's a question that resists an unequivocal answer. For one thing, news analytics firms sign agreements not to divulge the names of customers, which understandably don't want to expose investment strategies. (A number of larger firms reportedly have been using such algorithmic tools for a while.) Moreover, those analytic systems are diverse and changing fast. And apart from backtesting by vendors and rare testimonials, evidence of success or failure is largely circumstantial and sketchy.

To flourish, news analytics must boost quantifiable returns above the cost of installing and operating sophisticated systems. Leasing a news analytics system plus data feeds can cost \$5,000 to \$20,000 a month, depending on the features, frequency of data refresh and number of seats. That's particularly a challenge for newer, smaller firms.

And yet there are fans. Kevin Shea is confident that the payoff exceeds the cost. "If I can't get at least 3 to 5 percentage points of alpha from a factor, I'm not interested in looking at it," says Shea, a veteran of Cadence Capital Management, Batterymarch Financial Management and Invesco who launched Boston hedge fund Disciplined Alpha this year. As its name suggests, the firm adheres to a systematic investment strategy. Its algorithmic approach, developed by Los Angeles-based Alexandria, is rooted in bioinformatics, an information technology that emerged from genomics. Conceptually,

analyzing the genome sounds pretty mechanical; after all, DNA may be long, but it only has a four-letter code. But that code features deletions, mutations and “junk” sequences along its 3 billion base pairs, and its interactions with RNA and the assembly of proteins has proved to be extremely complex. Context and relationships matter, just as with financial information.

There are contextual elements in the way Alexandria “trains” algorithms to generate market insights. Rather than assigning positive, neutral or negative meanings to words in advance and imposing rules to classify sentiment, Alexandria analyzed 55,000 documents in one study deemed positive, neutral or negative by outside investment professionals and searched for deeper commonalities that supported assessments. When fed 5,000 new documents, the algorithm matched human assessments 91 percent of the time, says Shea.

Still, a 91 percent hit rate to one observer looks like a 9 percent miss rate to another. At high volumes that’s a lot of misses, though human assessments aren’t necessarily better. Moreover, the program processes tens of thousands of documents that otherwise might escape notice.

Without extensive stock screening affirmed by backtesting, formal risk models and sentiment signals, Shea says, Disciplined Alpha would find it difficult to generate satisfactory risk-controlled performance on a consistent basis. At least at Disciplined Alpha, this algorithm is a still-unfolding experiment.

Daniel Sandberg is also exploring the potential for algorithm-driven investment. Sandberg earned a Ph.D. in computational physical chemistry in 2012 from the University of Connecticut, then, like several of his peers, headed to finance. He joined the Legacy Foundation, an investment advisory firm in Charlottesville, Virginia. If algorithms can extract meaningful signals from scientific research, Sandberg saw no reason that they couldn’t work in finance. So he began to develop his own algorithms, working with venture-backed Boston start-up Quantopian, which provides tools like backtesters, data feeds, algorithm writing and a community of users. His first project: an algorithmic tool for implementing a sector rotation strategy.

For its part, Quantopian is trying to democratize algorithm development for smaller asset managers like Legacy and even for consumers. It claims to have the first algorithmic trading platform in a browser. And the company is planning to release a discount trading platform, meaning that algorithmic trading, and algorithm development, could be coming to retail investing.

SOME SKEPTICS CONTINUE TO FIND news analytics less than breathtaking. “Perhaps we’re behind the curve, but this isn’t a topic that we utilize much here,” says an equity research director at a global investment firm. New York–based Investment Technology Group provides many services and third-party research to hedge funds; in fact, one of its marketing slogans, “Decoding signal from noise,” could come from a big-data vendor. But ITG does not “traffic in news analytics,” says a spokesman. In late 2011, London’s Derwent Capital Markets launched a fund based on Twitter with great fanfare and celebrated a hefty 1.85 percent gain in the first month of operation. A year later, with returns far short of expectations, Derwent closed its Twitter-based fund.

Human judgment and oversight appear resilient to Christopher Cutler, a former chairman of the Alternative Investments Committee of the New York Society of Security Analysts. “Could this be a big game changer?” he asks. “I wouldn’t overestimate it. Too many things on the fundamental side of investing only humans can take a look at.” By way of illustration, Cutler cites a conversation between an analyst and a corporate executive that yields insight into the mispricing of products by rivals.

T. Rowe Price Group has long prospered with its traditional buy-and-hold investment strategies. “A process on auto-drive maybe will make some money, but in my mind it’s not the magic potion,” says Andrew Brooks, head of the firm’s U.S. equity trading. “Will 40,000 news releases put you into LinkedIn at \$6 and stay to \$200? I don’t think that happens by reading big data.” Portfolio insurance was hot in the mid-1980s, Brooks recalls. Then came the 1987 crash. What sent the market spinning? “Portfolio insurance,” he says.

News analytics can go only so far without humans, says Burke Lau, a Hong Kong–based market analyst at Macquarie Group, which endorses news analytics and, in partnership

with RavenPack, sells tools to global banking customers. Humans are needed to spot and correct false confidence in algorithmic models (and, perhaps, vice versa). A popular case in point: When new software at electronic trader Knight Capital Group malfunctioned in August 2012, unleashing a flood of buy orders on the exchanges, millions of dollars essentially vanished before a human — or humans — had to intervene.

Computers are suited to finding a specific data set and predicting an outcome. But when the unanticipated comes along, computers are ill equipped to respond without human intervention. For example, IBM Corp.'s chess-playing computer, Deep Blue, defeated Russian champion Garry Kasparov but was still unable to bluff or spot a bluff in poker — a game that resembles trading more than chess.

“In chess you can draw a circle around everything you need to know, whereas who knows what is going to affect the auto industry?” says Leslie Valiant, the T. Jefferson Coolidge Professor of Computer Science and Applied Mathematics at Harvard University and winner of the 2010 A.M. Turing Award for outstanding research in computing. Computers see a spectacular number of moves ahead in a chess game, but as of now, Valiant says, “we don’t know how to make them replicate common sense.”

Algorithms have more to learn before they can give investment advice with minimal human interference. But analysts who hope that judgment will always trump algorithms could face a rude surprise. Natural language processing, which allows algorithms to analyze everyday English, has made progress, as smartphone owners can testify. Reading a poker bluff requires sensitivity to semantic nuance. An IBM successor to Deep Blue named Watson, using multiple algorithms, famously defeated human champions on *Jeopardy!* partly by accurately parsing the game show’s wordplay.

“The advent of semantic search in financial markets stands to force a shift in the way financial professionals consume and analyze information and, most important, make money,” says Haris Husain, who heads a Thomson Reuters effort to develop intelligent search tools rooted in natural language.

And that may prove to be the rub for securities analysts, who have already suffered through more than a decade of wrenching change. With investment capital fueling start-ups and lots of shrewd engineering minds engaged in devising more-intuitive algorithms, securities analysts may face significant disruption. The best will survive and prosper; the rest may fall to the machine.

Christopher Steiner goes a lot further. Steiner, author of *Automate This: How Algorithms Came to Rule Our World*, is a former *Forbes* technology editor and a current Internet entrepreneur; he is bearish on the outlook not only for securities analysts but for professional portfolio managers. “In ten years I don’t see a whole lot of room for active managers of money,” he declares.

That may verge on the glibly apocalyptic, but there clearly are larger forces at work throughout the white-collar economy. Nobel laureate and Princeton University economist Paul Krugman echoed the warning in a [recent *New York Times* column](#) aimed at workers who think advanced degrees mean job security. “A much darker picture of the effects of technology on labor is emerging,” he wrote. “In this picture, highly educated workers are as likely as less educated workers to find themselves displaced and devalued.”

In such a scenario investment research may well require a smaller number of meta- or überanalysts, says Instinet managing director Joseph Mezrich, a quantitative analyst who convened a conference in May with a panel devoted to big data.

Still, these are the early days in news analytics, says MIT Sloan School of Management finance professor Andrew Lo, who directs the school’s Laboratory for Financial Engineering. Lo sees bullish prospects for algorithms that can extract data from the news. Put news analytics to work, he says, or else surrender innumerable opportunities to stay ahead of other investors when markets move. “No matter how much you try to predict market behavior,” Lo says, “it will always be the case that news volume will produce spikes in volatility.”

What we need to know is whether that news-driven volatility is just a passing fancy or a substantive development, a summer storm or climate change. If only we had an

algorithm to tell us that. • •

Read more about [asset management and trading and technology](#).