

# All the News that's Fit to Trade

*News has been moving financial markets since before Julius Reuter released his first pigeon. But what kind of competitive edge can machine-readable news give to algorithmic and automated traders? Chris Hall reports.*

Each new economic data release is like the start of a new ball game," says Paul Smilgius, founder and head trader at Sargis, a Chicago-based futures trading firm. "Because it's potentially the biggest market-moving event of the day, there's a lot of anticipation." To help ensure that Sargis makes the right call early in the game, Smilgius relies on a machine-readable economic news feed.

The low-latency feed – supplied by a specialist provider, Need To Know News – is programmed into Sargis' trading application which automatically fires off the appropriate trade(s) in response to the data contained in the specially-formatted economic data news story. "If US non-farm payrolls came out significantly below estimates, you might want to be a buyer of Eurodollars or short-term interest rates or you might want to be a seller of the S&P," Smilgius explains. "You need to be able to incorporate as much of the data into your trading strategy as possible because although payrolls may be down this month, the revision (to last month's figures) may go in the opposite direction."

Smilgius, a 15-year veteran in the Chicago Mercantile Exchange's trading pit, has been using machine-readable news to fuel Sargis' automated trading strategies since making the switch to screen-based trading four years ago. Starting out as a

market-maker at CME, LIFFE and Eurex US, Sargis is now largely focused on event-driven and momentum-based trades. "We've developed ideas on the screen that had worked in the pit, such as spread-trading and momentum trades, across a range of CME contracts," says Smilgius.

Because speed of execution is key to the success of Sargis' trading strategies, it was a logical step for Smilgius to feed machine-readable – also commonly referred to as digitised or elementised – news into the firm's trading applications with minimal latency. "We started out with another financial news provider, but soon realised we only wanted a specific piece of their coverage, i.e. the economic data releases," he says. "As well as needing the content being tailored to our needs, we wanted the fastest possible feed."



Paul Smilgius, Sargis

*"You need to be able to incorporate as much of the data into your trading strategy as possible ..."*



## Turning news into data

An increasing number of hedge funds, market-makers and prop shops looking for an edge over the competition are exploring the scope of machine-readable news to add an extra dimension to their trading strategies. "Everyone intuitively understands that news moves markets," says Alan Slomowitz, Director, Institutional Product Development, Dow Jones Content Technology Solutions. "Now traders have turned to more automated means of trading, they want as many different inputs as possible. They started by feeding raw market data into their trading models and algorithms, and now firms like ours have turned news into data that can also be very reliably integrated," he says.

Dow Jones is one of a growing number of firms – which now range from specialist US-based providers such as Need To Know News and Acquire Media to global information providers including Reuters and Thomson Financial – offering a range of products that enable financial markets clients to respond quickly to news events. The core products are news feeds for trading and archives for model-building,

but new 'sentiment analysis' tools are supported increasingly sophisticated use of machine-readable news. Low-latency real-time news feeds contain textual information that has been formatted, tagged and identified so that an automated trading model can use the content to take particular trading actions, either on its own or in combination with other data. Archives contain many millions of machine-readable news items from up to 20 years ago that can be used in conjunction with tick data to identify and exploit historical trading patterns.

Providers make news stories machine readable first by putting key elements in a consistent XML format and second by adding meta-data tagging to identify companies, industries, subjects, geographies and other facts. For Thomson Financial, whose archive draws from third-party and acquired news sources, this has meant applying a consistent format and clean ontology across 50 million news stories going back ten years. "Both automated trading strategies and execution algorithms have traditionally relied on only real-time pricing information, lacking the ability to recognise market-moving events in the news," says Ryan Terpstra, Senior Strategist, Corporate Strategy Group, Thomson Financial. "As a result, people are trying to build news-intelligent models and algorithms to incorporate textual data."

## When bad news is good news

Use of real-time machine-readable news feeds that enable automated trading models to respond quickly to scheduled news items such as economic data releases has rapidly expanded to encompass a wider range of predicted and unpredicted stories. While it is a small step from trading off breaking economic data to responding automatically to scheduled corporate earnings news, many firms are using machine-readable news feeds for risk management rather than alpha-generation purposes at present.

"A bank, mutual fund, or hedge fund that is making markets or holding a significant long-term position is susceptible to short-term market moves if some

economic numbers come out that are way out of line with expectations,” says Justin Nolan, Director of Marketing, Need To Know News. “They can use this information as a hedge or insurance policy and cover themselves for a short-term move during a long-term position hold.”



Ryan Terpstra, Thomson Financial

*“... automated trading strategies and execution algorithms have traditionally relied on only real-time pricing information, ...”*

US-based Acquire Media delivers a low-latency XML-based corporate announcement and economic news feed to asset managers, hedge funds and professional investors that combines the output of all six North American electronic press release wires, a rigorously-maintained calendar of scheduled announcements by listed firms and news agency reports.

“A market-maker will have many active bids and offers out there, and doesn’t want to get hit when unexpected news comes out, such as a special dividend on a particular stock,” says Don Rigg, Head of Sales, Acquire Media. “A buy-side firm that’s feeding a large order into the market will want to stop an execution algorithm automatically if the firm is suddenly hit with a major law suit.” Because some stories almost always impact a firm’s share price (e.g. law suits, downgrades) while others (most product announcements) may not, Acquire Media allows users to specify category of news that can stop a trade.

As well as alerting a busy trader to switch off an execution algorithm or automated trading model to digest the impact of a news announcement, machine-readable news feeds can effect an instant switch in strategy. “Research suggests that news flow is a good indicator of volume and volatility,” says Richard Brown, Global Business Manager, Reuters NewsScope. “If you’re executing a USD 1 million stock purchase order using a VWAP algo and news comes out on that stock, you might want to switch to a more participatory algorithm in order to take advantage of a potential volume spike,” he says. “The line between alpha algorithms and execution algorithms is getting blurred.”

### All inputs considered

At the same time, quantitative model-builders are using machine readable archives to improve automated trading strategies and uncover new investment ideas. Steve Wilcockson, Financial Applications Manager, The MathWorks, producer of the MATLAB model-testing environment, says machine-readable archives can have significant benefits to backtesting. If, for example, a news archive can demonstrate that the relationship between two equities has been disrupted by a particular type of news event in the same way over time (e.g. the correlation widens by 20-30 ticks), a pairs trading arbitrage model can be made more effective by adapting it to take account of this news event on future price movements. “When working with a pairs arbitrage model that trades intraday, the ability to explore what’s causing the mean reversion, can be interesting,” says Wilcockson. “If you’re trading one company against another and one of them has ongoing negative news about a production facility, for example, we might want to regress the performance of the stock with respect to this news and therefore refine our backtest.”

New strategies are also being developed through analysis of tick data alongside machine-readable news. Armando Gonzalez, President and CEO of RavenPack International, a firm that specialises in developing news analytics for use in algorithmic and automated trading, says hedge funds in particular are hungry for new data inputs into trading models. “If you know that 25 CEOs of public firms quit their jobs over a five-year period, you might do correlation analysis of the stock price performance for the next hour after the news was released,” says Gonzalez. “If you find that 75 per cent of the time, the price went down within a certain percentage range in a given period, you can use that to programme models.” ▶

But use of machine-readable news items in model-building may only be the first step of an expansion into a much wider range of data inputs. If models can be built to react to news items, how much more effective could they be by harnessing other information sources? Reuters' Brown says firms are looking at "not so obvious" correlations between news events and market movements. "Typically, if a hurricane is upgraded from 3 to 4 in the Gulf of Mexico, that's going to cause problems for the stocks of oil firms drilling there. But you also need to integrate that news with other data sources that identify which firms have rigs and where. If you know there's a strike at a chip manufacturing plant in China, you also need to know whether the firm is supplying Dell or IBM," says Brown. "The advantage is measured not in milliseconds, but perhaps days, weeks and months if you identify an investment idea ahead of the market. Capturing alpha depends on mastering all those data sources."

To an extent, machine-readable news products are enabling the faster end of the quant-based automated trading market to incorporate strategies more commonly associated with traditional asset managers. The failure of some models to account for the impact of the sub-prime mortgage crisis has piqued hedge funds' interest in news analytics according to Ravenpack's Gonzalez. "They found their models lacked a sense of awareness of the world beyond fundamentals and technical analysis," he says. "Previously, some funds' models had been fully dependent on analysis of tick by tick data and fundamental company data. Backtesting would reveal discrepancies between fundamentals and tick data that could not be explained unless one allowed for an element of interpretation by analysts and the media," says Gonzalez.

The MathWorks' Wilcockson agrees that firms are looking to widen their frame of reference. "Over the last two to three years, a number of higher frequency systematic desks that have been traditionally very market data and returns data-driven have been embedding fundamentals data (i.e. price to earnings ratios and other balance sheet data) on top of their models," he says. "And now they're thinking about incorporating news reports, e.g. comment from analysts, which is something that longer-term desks have been playing around with for a number of years. Machine-readable news helps that process in this time-constrained high frequency environment."

### Sentimental education

One of the main ways in which traders are looking to incorporate a wider range of news inputs into automated trading models is through sentiment analysis tools. These 'sentiment engines' interpret individual news items at high speed and give them a numerical rating which the trading model can respond to. "Most interest is among firms that trade frequently as previously it was too labour-intensive to take the impact of news on stat arb-type trading models," says Wilcockson. "The speed at which sentiment engines can now interpret responses to news changes the picture."

Reuters NewsScope Sentiment Engine, developed in partnership with Infonic, an AIM-listed linguistics technology firm, enables users to track the sentiment on a particular company, benchmark it against peers, or aggregate the sentiment for a portfolio, or even an entire market, in real time. Dow Jones News Analytics, which uses proprietary algorithms to tag stories as positive, negative or neutral, was co-designed by RavenPack International. Working with clients on an individual basis, Ravenpack offers three methods of analysing and categorising news stories.

For the 'traditional' method, Ravenpack used feedback from a group of analysts to create a list of thousands of positive, negative or neutral words or phrases commonly used in the financial markets that, when read in a news story, would tend to make the analysts consider buying, selling or holding a stock. In practice, a rules-based library is used to scan the text and score the story, first on whether it is positive, negative or neutral and then on the degree of the sentiment, based on the number of positive, neutral or negative phrases identified. The 'expert consensus' method draws on the collective opinion of Ravenpack's pool of analysts on whether a sample of 10,000 stories would have a positive, negative or neutral impact on market sentiment. Ravenpack developed a method that replicates the analysts' consensus view on new stories. "Clients can multiply scores generated by the expert consensus engine by a negative value in order to take a contrarian view of the market," says Gonzalez. The third method, 'market response', is based upon how markets reacted to similar types of news story in the past. "We might find that every time a particular central bank governor uses a certain combination of phrases to describe the economic outlook, there's a strong positive reaction in the equity markets," adds Gonzalez. ▶

The ability to customise and tweak sentiment engines could be crucial to wider uptake says The MathWorks' Wilcockson. "If sentiment engines effectively turn a statement into a number based on how many times specific words or phrases occur, some users will want to look very closely at how this linguistic analysis has derived an output," he says. However, Wilcockson asserts that sentiment analysis is likely to attract firms that trade heavily in the reporting season, using models that respond to actual performance (balance sheet data) against market expectations as well as the market's subsequent reaction to that performance. "Another sector that could benefit would be sell-side firms that trade on the volatility that occurs during reporting periods. At the moment, a lot of their analysis looks primarily at the returns level, but they will be interested in sentiment analysis as a cause of the price changes during this period," says Wilcockson.

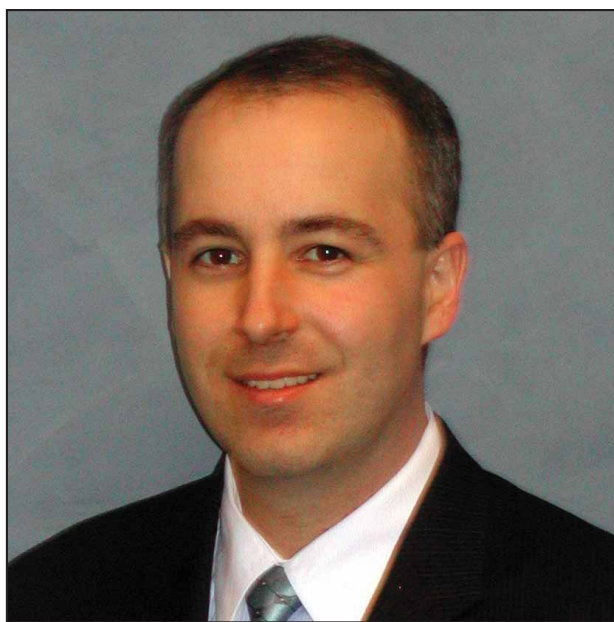
In addition to buy and sell signals for individual stocks, Brown suggests sentiment analysis could also be used for portfolio reallocation purposes, i.e. based on a comparison of the overall sentiment of news stories about NASDAQ-listed stocks versus those traded on the LSE, or emerging markets economies versus European stocks. "A client might want to incorporate sentiment scores into an automated trading strategy by writing rules that trigger a buy or sell order once a certain stock exceeds a five-day moving average, either based purely on a particular sentiment indicator or in combination with other data," he says.

### Tomorrow's headlines

The development of tools to exploit the trading potential of machine-readable news continues apace. Thomson Financial, for example, provides an integrated offering which not only provides a news feed and archive, but also the tools to analyse them. Targeted at the high-end quant market, Thomson Quantitative Analytics normalises data sets from different financial market data and news vendors and provides time series manipulation tools to support backtesting.

Reuters NewsScope Event Indices, to be launched early Q2 2008, measures the market impact of different categories of news. Initially focused on the FX market, the tool can be used to predict volatility between currency pairs. The product measures the quantity of news within 45 topic-specific indices (e.g. political violence, natural disasters, macro-economics) over a set period. The indices are

mapped to 16 major currency pairs so that, as the quantity of relative news volumes within the indices reach certain levels, they serve as a predictor of volatility for the specific currency pairs. "Different types of news affect different currencies in different ways. The dollar-yen is affected by a different mix of the topic-specific indices than other pairs. We've found a volatility indicator, but combined with a client's proprietary research, it could help identify direction and magnitude," says Brown.



Richard Brown, Reuters

*"The advantage is measured not in milliseconds, but perhaps days, weeks and months if you identify an investment idea ahead of the market."*

Nevertheless, Ravenpack's Gonzalez believes the machine-readable news industry is barely beyond start-up phase, with only a fraction of news organisations' total daily output capable of being incorporated into trading models. "In 2008-09, I expect a high level of growth, including use of sentiment analysis on analyst commentaries and incorporation of company earnings announcements directly into quantitative models," he says. In the foreseeable future, a broader range of risks, from terrorism to pandemics, will be quantified for trading and risk management purposes. "The end will be when all elements of a news feed can be used to trigger a meaningful trade – then we'll have true automated trading." 