

WHITEPAPER

Gnip, Inc.
www.gnip.com
888.777.7405
trading@gnip.com
@gnip

SOCIAL MEDIA AND MARKETS: The New Frontier



Source: Gnip, Inc.

For the first time in history, access to the observations, wisdom and emotional reactions of millions of people globally is available in realtime. Social media data represents a collective barometer of thoughts and ideas touching every aspect of the world. As social platforms increasingly become a primary means of communication for our age, asset managers, equity analysts and high frequency traders are incorporating leading indicator data from these platforms into investment decisions as a means to create alpha.



CONTENTS

Abstract	1
Social Media Megatrends	1
These Financial Times: Faster Than Ever	1
Why Social Media Matters	2
Speed	2
Amplification	3
The View From The Ivory Tower	3
Indiana University	4
Pace University	4
Munich Study	5
How Social Media Factors In Trading	6
What's Next: Leaping In	7
Time To Change	8

Abstract

This paper discusses the growing use of social media among asset managers, equity analysts and high frequency traders, and specifically how social data streams can help make trading decisions. The paper documents the rapid growth of social media and defines how social media activity can be measured and used by investment professionals. It also outlines examples from academic studies and the stock market to provide investors with a list of important considerations when implementing a social media trading strategy.

Social Media Megatrends

Social media is growing exponentially. Facebook has exploded from one bored college sophomore to 800 million users in nearly 200 countries in just seven yearsⁱ; Twitter users generate 250 million new Tweets each dayⁱⁱ; and LinkedIn now has 135 million peopleⁱⁱⁱ trying to make connections, find jobs and lead discussions. Looked at in one way, social media is a vital pipeline of thoughts, words and decisions between two people, 20 people or a larger group of people sharing, connecting and interacting as never before. Looked at another way, it's the collective weight and barometer of thoughts and ideas about every aspect of the world – a collective pulse of observations, wisdom and emotional reactions for asset managers, equity analysts and traders to analyze and quantify.

These Financial Times: Faster Than Ever

The financial world has always been one of fortunes earned and squandered. Those who lived through the famous Black Monday of 1987, when the Dow Jones Industrial Index dropped more than 22 percent in a single day^{iv}, can testify that sharp swings in the market are nothing new. But since the credit crunch began in August 2007, days of high volatility have become the new norm. What is up one day can come crashing down the next. Over half the trading days in October 2011 had swings of greater than 2 percent in the S&P 500.

The ability to understand the forces shaping the market – allowing one to predict direction – has

*Looked at another way,
it's the collective weight and
barometer of thoughts and
ideas about every aspect of
the world – a collective pulse
of observations, wisdom and
emotional reactions for asset
managers, equity analysts
and traders to analyze
and quantify*

become an increasingly elusive skill. Traditional models, too stilted for our interconnected world, require modification – new data sources matched to the increasing speed of the world.

Social media is one of the data sources filling this gap. Equity predictions are appearing online with incredible speed, through platforms like StockTwits, featuring real-time conversations on market conditions and specific securities. Millions of people are creating a new conventional wisdom, highlighting the voices and trends that will move the markets.

Why Social Media Matters

Moving from the abstract to actual examples, first we need to discuss the two primary ways social media activity can be measured in aggregate: speed and amplification.

Speed

Social media is the closest humanity has come to creating a way to instantaneously monitor the pulse of the world and observe the stream of human consciousness. There are no traditional media processes to hinder and slow down the receipt of information. The death of Osama bin Laden first entered the public sphere through a Tweet by a former Defense Department official^v briefed on the forthcoming announcement.



Source: Twitter. <http://twitter.com/#!/keithurbahn/status/6487790624886784>.

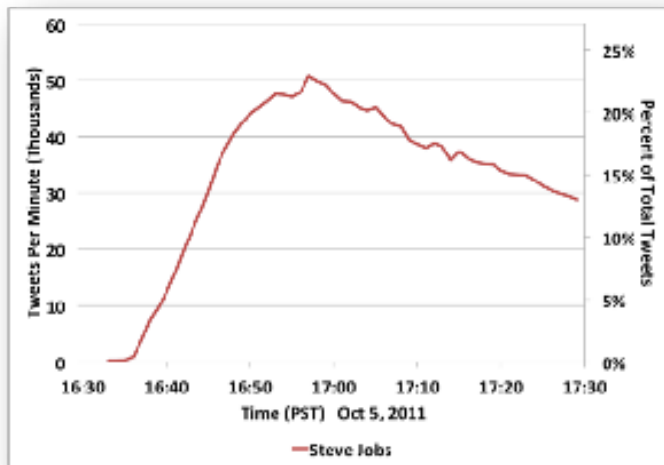
DataMinr, a social media monitoring tool, was able to spot this with just 19 Tweets on the subject. The company then issued a signal to their clients^v, alerting them to this important piece of information; .

*Millions of people are creating
a new conventional wisdom,
highlighting the voices and
trends that will move the
markets*

it would be over 20 minutes before that story appeared on traditional news sites. Access to a data stream that can beat traditional media sources by over 20 minutes requires no explanation as to its value for traders and investors. Speed matters.

Amplification

Amplification speaks to the ability of social media as a “crowd-sourced megaphone.” The propensity of users to like, share and retweet content from other users gives those consuming social media data an extremely easy mechanism to measure what content is most important to the world – and to compare that content against other content in real time. A prime example is the recent passing of Apple CEO Steve Jobs. Word of his death spread rapidly, and thirty minutes after the news broke 50,000 people were Tweeting about him each minute^{vi}. This was such an important story it occupied almost 25 percent of all Tweets at that time.



Source: Gnip Blog, “Steve Jobs Rest in Peace,” <http://blog.gnip.com/steve-jobs-rest-in-peace/>.

This data can be used by investment professionals to determine the relative weight or broad market interest in a particular story. By capturing a realtime snapshot of what the market considers important – and what it doesn’t – asset managers and traders are able to add an important factor to their existing algorithms in the quest to create alpha.

The View from the Ivory Tower

The utility of social media in financial decisions has

Access to a data stream that can beat traditional media sources by over 20 minutes requires no explanation as to its value for traders and investors

By capturing a realtime snapshot of what the market considers important – and what it doesn’t – asset managers and traders are able to add an important factor to their existing algorithms in the quest to create alpha

attracted the attention of economics professors around the world. While research is ongoing, the initial findings of three studies have been promising:

Indiana University

An Indiana University study^{vii} looked at more than nine million Tweets posted between March and December of 2008 in an effort to see if small-form communication could be used to predict the market. Professor Johan Bollen, along with two graduate assistants, sorted the Tweets into categories – calm, alert, sure, vital, kind and happy – and compared them to the market. The researchers found that the calmness index is a good predictor of whether the Dow Jones Industrial Average (DJIA) goes up or down between two and six days later. The calmness of Twitter on any given today could predict moves in the Dow over the next week with 87.6 percent accuracy.

In a *Technology Review*^{viii} article examining the study, several important caveats are discussed, most importantly the lack of geographic clarity as to Tweet origin. As the Tweets were not filtered by geography, it is unclear if the study represents a true cross-section of the United States at any given time. But the results are intriguing enough to have inspired multiple funds to pursue a trading strategy incorporating this approach.

Pace University

Arthur J. O'Connor from New York's Pace University looked at three Blue Chip brands^{ix} – Starbucks, Coca Cola and Nike – over ten months on social media. He tracked the number of Facebook fans, Twitter followers and YouTube views to measure social media popularity.

He tracked this data against daily stock price movements and an index of consumer stocks. The study also looked at keyword search trend data to determine changes in search volume. Even though the performance of the stocks varied dramatically – Coke declined by about 6 percent, while Starbucks rose 29 percent and Nike was up 14 percent – a correlation was found between daily popularity and stock price. The correlation remained when

The calmness of Twitter on any given today could predict moves in the Dow over the next week with 87.6 percent accuracy

researchers introduced a 10 and 30-day lag into the study, suggesting social media popularity may be a leading indicator of stock price performance.

While the study showed that social media's "word-of-mouth" power is a fast way to measure consumer sentiment, additional research is needed to see if social media can move beyond consumer brands with large Facebook followings.

Munich Study

A final study from the TUM School of Management in Munich^x examines two primary questions: (1) how the content or information contained within stock microblogs reflect financial market developments and (2) if microblogging forums are an efficient platform to weigh and aggregate information. To conduct this study, researchers focused on conversation explicitly concerned with the market, collecting nearly 250,000 tweets over a 6 month period (Jan 1 - June 30, 2010) that all contained a \$TICKER symbol for a company in the S&P 100. Their findings were interesting for what signal they found, as well as some of the caveats they raised.

The first finding concerned microblogged information in its relation to market efficiency. The research found that stock microblogs do contain valuable information that had not been priced-in yet – data not incorporated into current market indicators. The market quickly corrected for that information, however, making it difficult to exploit given normal transactional costs. But as the study points out, more sophisticated algorithms might be able to take advantage of this information arbitrage.

The second interesting finding relates to the predictive correlation between microblog message volume and following trading volume for a specific stock. This correlation is clear even with the inclusion of numerous control variables and correlates messages from one to two days prior with current day trading volume.

Finally, the authors found that microblogging mechanisms like retweets and follower counts are closely related to an efficient diffusion of information. The study finds that users who provide

*But as the study points out,
more sophisticated algorithms
might be able to take
advantage of this
information arbitrage*

above average advice in the investment arena do indeed have more retweets and a greater follower count. Interestingly, however, analysis of individual messages shows that higher quality pieces of information are not, themselves, retweeted more frequently than others. So consideration of user data (like follower count) can help weigh the value placed on specific information for a specific stock.

How Social Media Factors in Trading

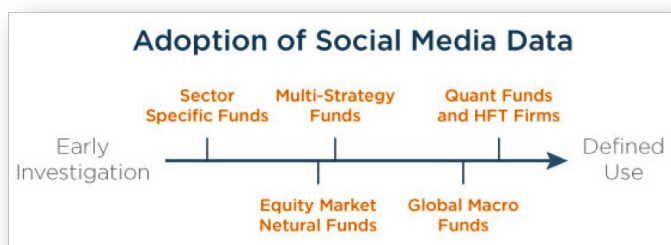
As investors innovate in how they leverage social media as a data stream, several trading use-cases have surfaced. First and foremost is the use of social media for sentiment analysis – a measure of how people feel. Investment funds who are using the social media “firehose” are applying complex algorithms to measure an array of sentiment trends. This can be as broad as considering the overall ‘market sentiment’ (what is the mood of the entire consumer marketplace?) or as specific as examining positive and negative emotions or opinions associated with specific brands or equities. What’s fascinating is their ability to do this in real time, given the low latency of specific social media streams like Twitter and StockTwits.

A second use-case focuses on social media as a ‘breaking news’ stream. The DataMinr example discussed previously (regarding the death of Osama bin Laden breaking on Twitter 20 minutes before traditional news channels) is a clear example of Twitter as a newswire.

Other use cases often discussed include macroeconomic trend analysis (data on consumer retail spending, the housing market, job searches, etc.), and currency exchange (geographic risk alerts, country-specific analysis, etc.).

As more funds and traders gain access to this data, we continue to see greater innovation in use of the data as a predictive factor. Use in trading algorithms is growing, but the adoption rates vary depending on the type of fund and their strategy. The chart found at the top of the next page outlines adoption of social media data by investment strategy.

What's fascinating is their ability to do this in real time, given the low latency of specific social media streams like Twitter and StockTwits



Source: Gnip, Inc.

What's Next: Leaping In

When a firm decides to use social media in a trading strategy, there are several key factors to consider in selecting a data provider. The three most important considerations when choosing a provider are where the company sources its data, their infrastructure and their understanding of the breadth of social media sources and how they relate to a fund's intended use case.

Users need to determine the legality and providence of the data they are to receive and there are certain tradeoffs to consider, such as potential volume of data versus long-term reliability of that data. Some data providers are driven through publisher agreements (both exclusive and non-exclusive) or public API data feeds, and only collect information that they are given. These companies refuse to use private information. Other data providers "crawl" the web to source their data. Sources that crawl may be able to deliver a larger volume of content from blogs, boards and forums, as well as social media sites that only provide limited API access (like Facebook). A challenge in using them, however, lies in the unreliability of their service. Crawlers have a greater potential to be shut down or blocked than publisher-agreement providers, who may be more limited in volume by their adherence to public data.

Of course, regardless of how data is collected, it's useless if a company can't access it. When choosing a provider, ask the right questions - How reliable are their connections? What's the latency in the system? What does API access look like? How long will implementation take? Different firms provide the data in different formats and latency. Delving deeper into infrastructure questions prior to commitment- or

The three most important considerations when choosing a provider are where the company sources its data, their infrastructure and their understanding of the breadth of social media sources and how they relate to a fund's intended use case

trialing the service for compatibility - is very common.

For simplicity we've focused on several key social networks in this paper, but breadth of coverage is another consideration in choosing a provider. Academic and investor interest to date has focused on Twitter, but social media is as dynamic as the constantly shifting opinions featured on it. Google+ is an emerging competitor to Facebook as a user-profile and connection platform. StockTwits is a niche player catering to active investors and traders looking for real-time discussion. Foursquare and location-based services have the potential to provide important information about traffic and trends to specific retail locations. Blogs and forums are valuable sources of deep insight. Companies need to probe potential data providers both on the current offerings and how they decide when to offer new products. Maintaining the information edge requires constant work. A data provider should be able to break down the differences in platform and data source in relation to the intended use of the data by the inquiring fund or trader.

Time to Change

Not everyone wants or needs to use social media data as part of the investment process. But as social media garners greater mind share, accessing Twitter data on a company may be as common as checking a stock quote.

The challenge of incorporating social media into traditional research and review strategies is rapidly being overcome as new products and services enter the marketplace. Whether it is aggregating high volume, low-latency new media, sifting through social data across many sites, or providing real-time data streams for trend analysis, hedge funds are driving innovation as service providers seek to give them access to leading indicator data that will help them create alpha for their investors.

Companies need to probe potential data providers both on the current offerings and how they decide when to offer new products

REFERENCES

- i Facebook. <http://www.facebook.com/press/info.php?statistics>.
- ii Mashable. <http://mashable.com/2011/10/17/twitter-costolo-stats>.
- iii LinkedIn. <http://press.linkedin.com/about>.
- iv Daily Markets. "Great Stock Market Clashes Black Monday," <http://www.dailymarkets.com/economy/2009/08/20/great-stock-market-crashes-black-monday-in-1987/>.
- v All Things D. "Meet DataMinr One of Twitter's Favorite Start-Ups," <http://allthingsd.com/20111005/meet-dataminr-one-of-twitters-favorite-ecosystem-start-ups/>.
- vi Gnip Blog. "Steve Jobs Rest in Peace," <http://blog.gnip.com/steve-jobs-rest-in-peace/>.
- vii Cornell University Library, "Twitter mood predicts the stock market" <http://arxiv.org/abs/1010.3003>.
- viii Technology Review. "Twitter Mood Predicts The Stock Market," <http://www.technologyreview.com/blog/arxiv/25900>.
- ix VentureBeat. "Social media popularity can predict stock prices," <http://venturebeat.com/2011/03/17/study-social-media-popularity-can-predict-stock-prices/>.
- x VentureBeat. "Twitter predicts performance of individual stocks," <http://venturebeat.com/2011/04/06/twitter-predicts-performance-of-individual-stocks/>.

WHITEPAPER

Gnip, Inc.
www.gnip.com
888.777.7405
trading@gnip.com
@gnip

About Gnip

Founded in 2008, Gnip developed the first enterprise-grade platform to reliably deliver the billions of realtime public social media activities that power our customers' businesses every day. Gnip provides access to more than 100 realtime social media data streams from dozens of social media sources, including Twitter, Facebook, YouTube, WordPress and more. Today, we deliver our data, ready-to-use, in the protocol of each customer's choice. We apply data enrichments including URL expansion, language detection, Klout Scores and more to augment the raw data, making it even more valuable for our customers. For more information, call 888.777.7405 or visit <http://gnip.com>.

About The Author

Seth McGuire serves as the Director of Asset Management and Financial Technology for Gnip.

