



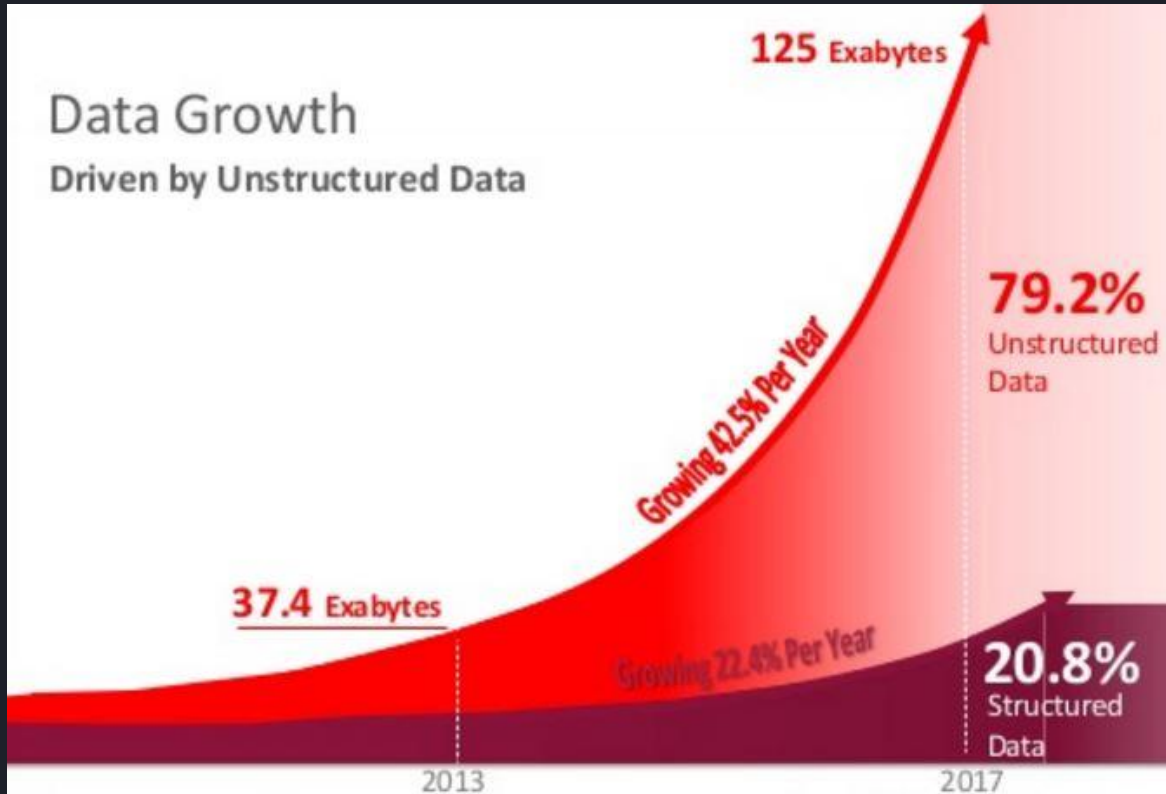
Alternative Data in Finance

Bringing Semantic Intelligence to Financial Markets



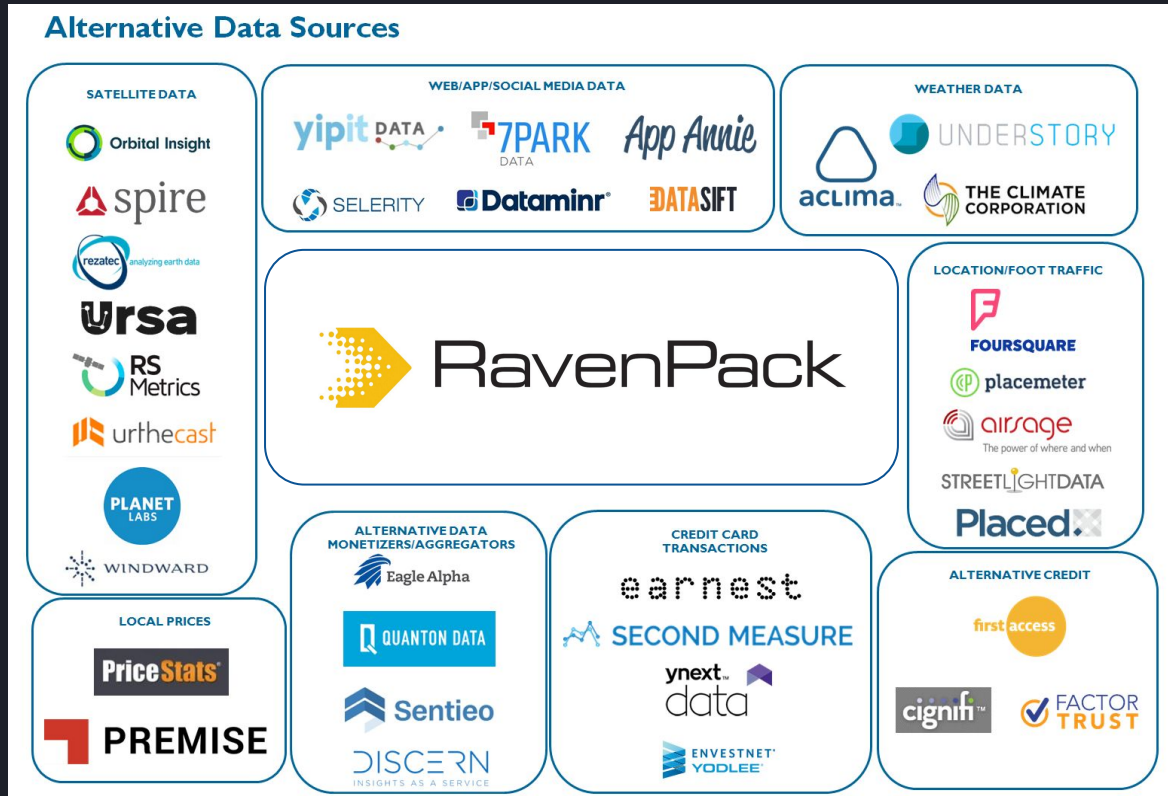
RavenPack

Data is growing at an incredible speed



- **90%** of all data that exist today has been generated over the last 2 years.
- Nearly 80% comes as 'hard-to-consume' **unstructured content**.
- Offers an **incredible opportunity** for investors to identify new alpha sources.

A new wave of data sources...

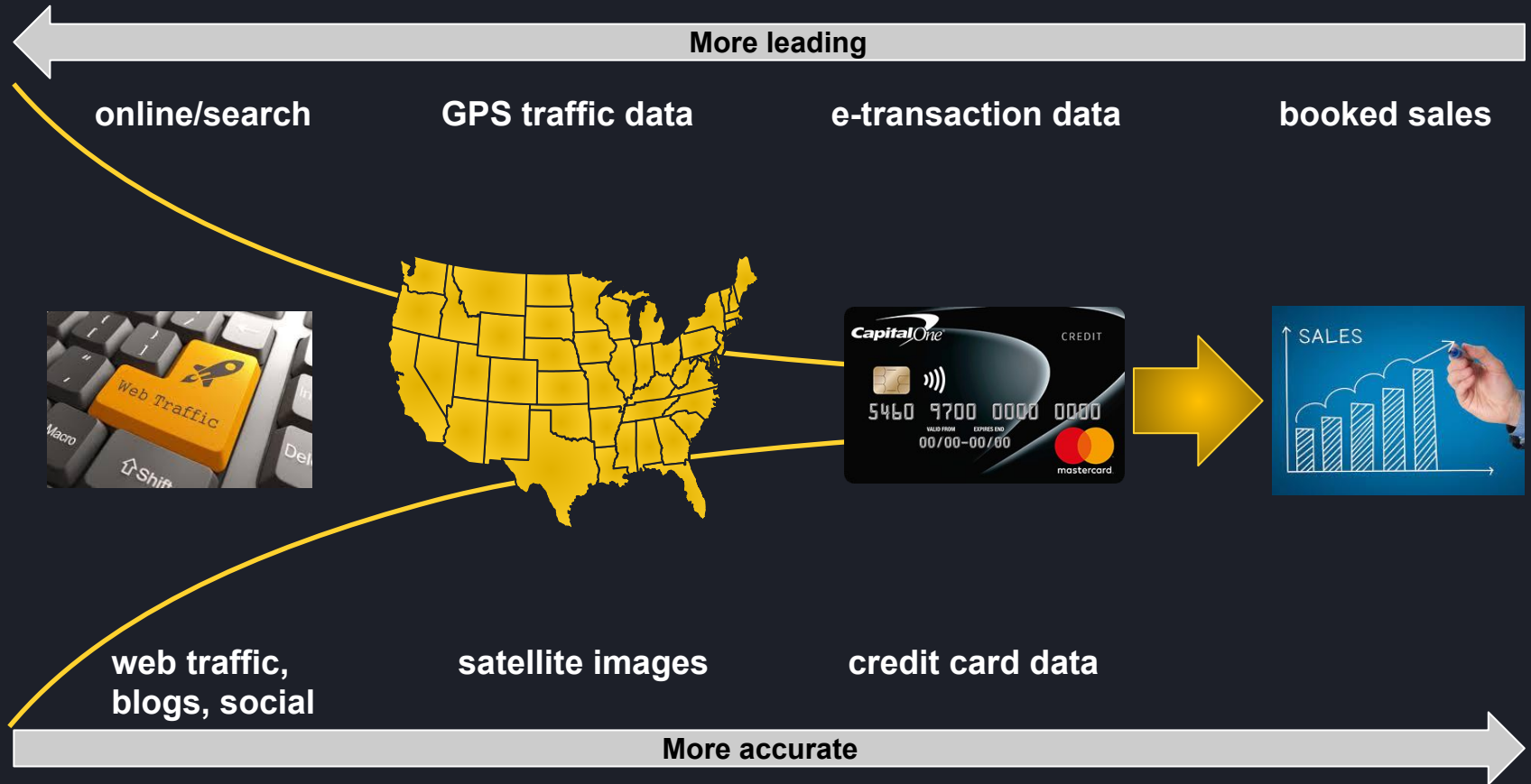


- **Thousands of data sources** have become available.
- **Early adopters** have had a real edge by hiring dedicated data hunters.
- **The market is maturing** with more sell-side research becoming available.
- Data hunting is becoming **less of a differentiator.**

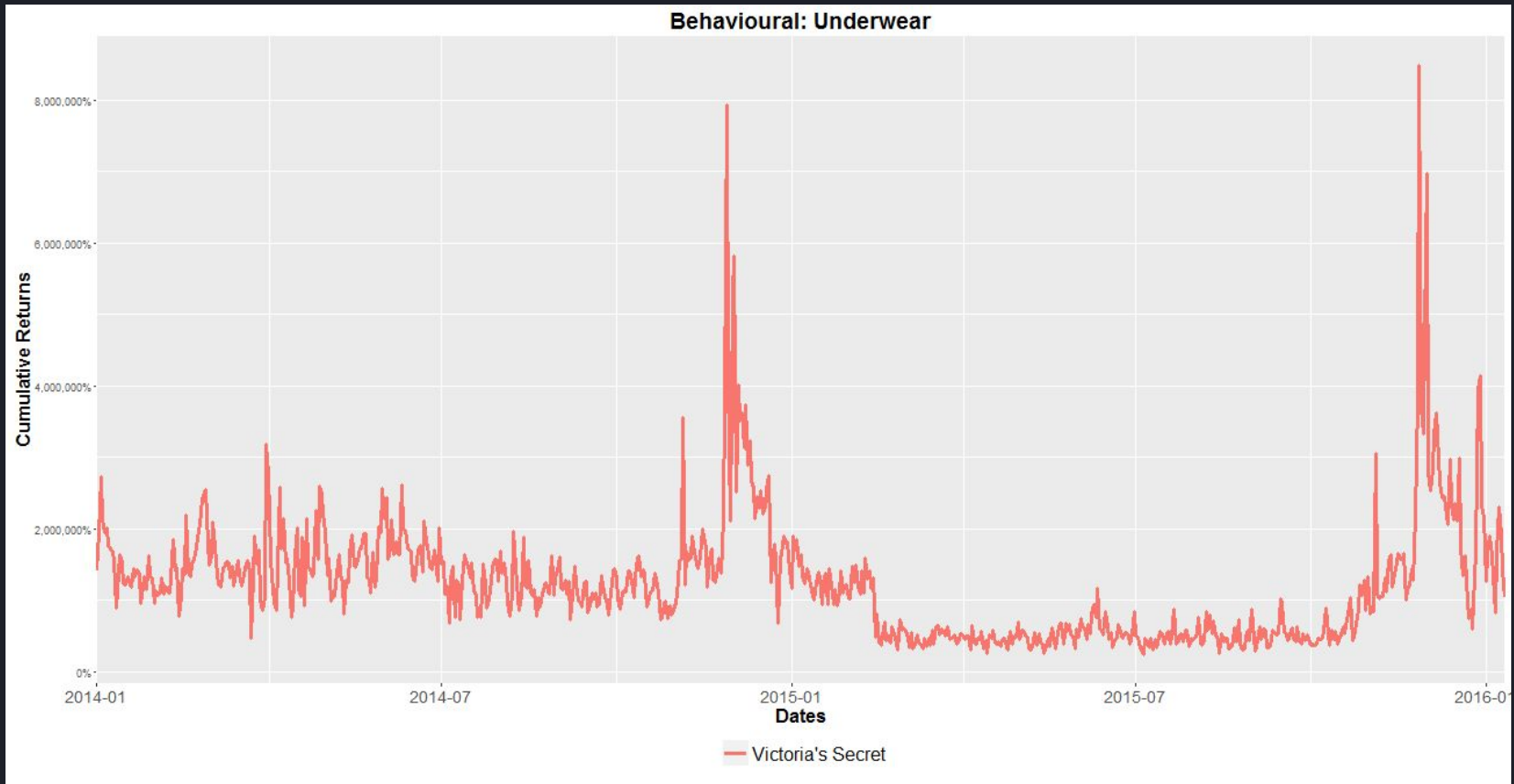


Make Sure **Data is Legal**

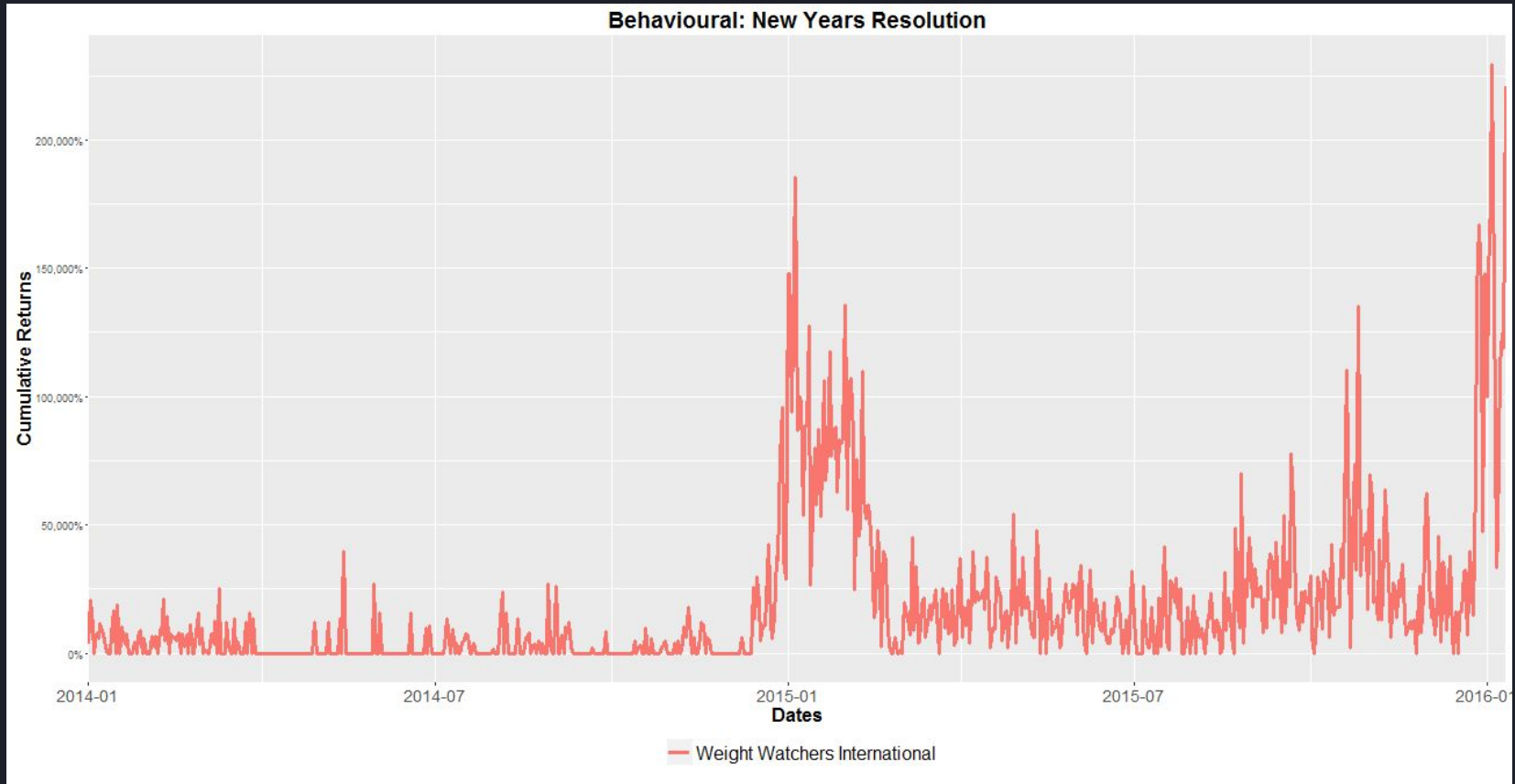
Why **alternative data**?...intuitive



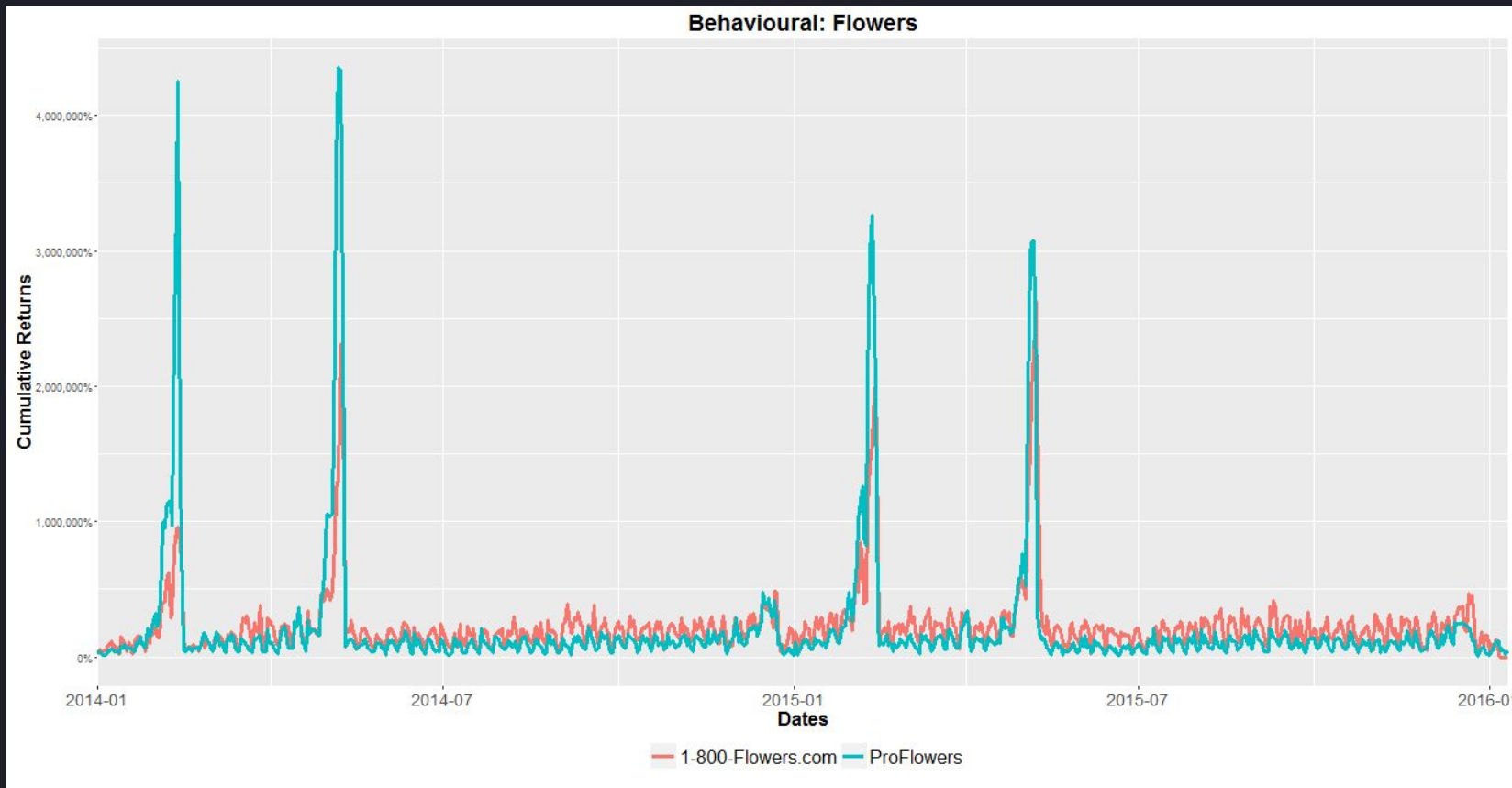
Behavioural Patterns: **Sexy Lingerie**...who's buying?



Behavioural Patterns: **New Year's Resolution...**



Behavioural Patterns: **Flowers** always help...

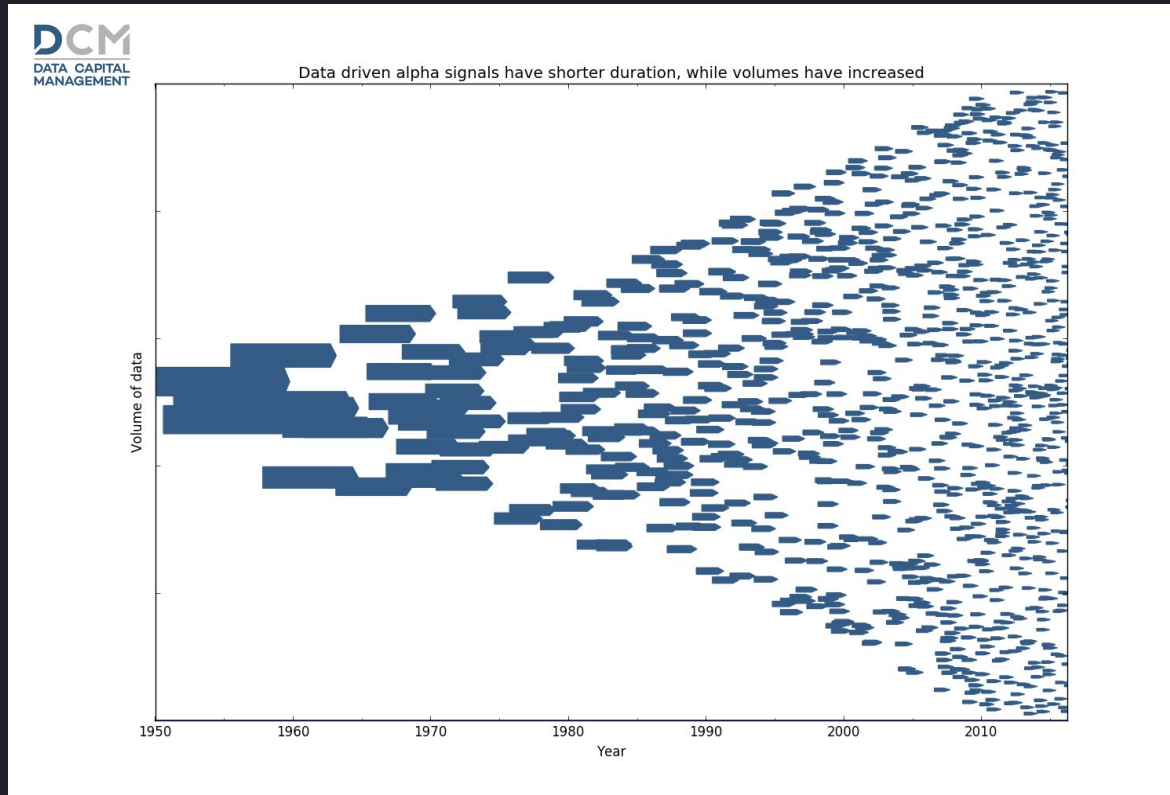


Proprietary vs. Public **Information**



- The edge is found in **efficient processing** not in being the only one having the data.
- **Old school thinking...** unless you're Google, Facebook, Apple, Amazon, or Microsoft.
- Many **correlated datasets** are available.
- Alternative data is in most cases about **nowcasting of fundamental** data.








The **alpha landscape** has changed...



Source: Data Capital Management (DCM)

- Over time, data driven alpha signals have **shorter duration**, while the number of alpha sources have **increased**.
- As a result, investors need to **consume more data** to create equally scalable strategies.
- There is **pressure on cost**, since each individual alpha contains less marginal value.
- More **pressure on infrastructure**: storage, computing, and research.

The Battle: centralized vs. decentralized quant alpha

Crowd-sourced Alpha	Talent Recruitment
 <p>NUMERAI</p> <p>A hedge fund built by a global community of anonymous data scientists</p>	 <p>kaggle competitions</p>  <p>TWO SIGMA</p> <p>FINANCIAL MODELING CHALLENGE</p>
	
 <p>Quantopian</p>	 <p>WINTON</p>

- **Big Data is a 'numbers game'**. Borrowing from Machine Learning, investors are looking for 1000s of alphas (weak learners) - taking an **ensemble learning approach**.
- **Talent Wars**: crowd-sourced alpha platforms (**decentralized**) vs. Alpha competitions (**centralized**).
- The **crowd-sourced alpha** approach sounds appealing, but is **still unproven** and suffers from various issues.
 - Access to premium content is hard (encryption)
 - Domain expertise/data specialization
 - Managing IP & brain drain



We make
unstructured content actionable
for financial professionals

Adding Structure



Codify

Named entity recognition, topic categorization, and temporal classification



Relevance Tagging

Quantify importance, rank, and connections



Novelty Detection

Identify what is new, original, or unusual



Sentiment Analysis

Determine attitude, opinion, and impact

Documented **Third-Party Use Cases...**

J.P.Morgan

Creates global macro strategies focusing on **Equity Index, Forex, and Sovereign Bond trading**

Duke
UNIVERSITY

Creates supply and demand side sentiment indexes to **trade crude oil**

TrendPointers
Business Trends & Research

Creates sentiment indexes for **Equity Index market timing**

THE THALESIANS
Cross Asset / Quant Strategy

Creates global macro strategies focusing on **FX Carry and Sovereign Bond trading**

NOMURA

Shorts stocks based on negative sentiment (Sell on the news)

Deutsche
Bank

Enhances a **pairs-trading strategy** using an abnormal news volume and sentiment overlay

WOLFE
RESEARCH

Applies various **machine learning algos for portfolio construction** using news sentiment factors

MACQUARIE

Pioneers in alternative data within **factor investing and event-driven strategies**

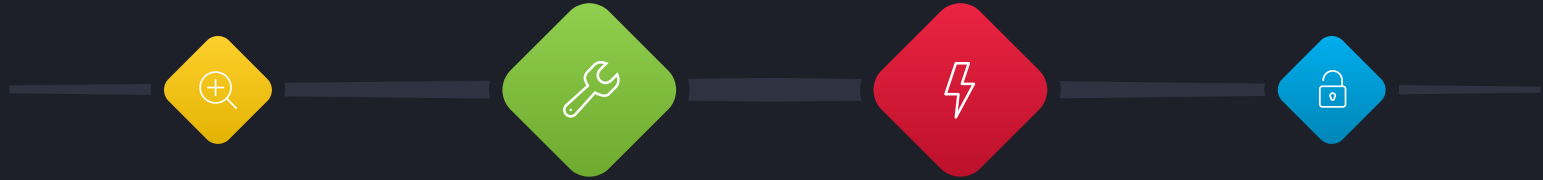
THE ALLIANCE
INSEAD **Wharton**
UNIVERSITY OF PENNSYLVANIA

Considers news sentiment for **intraday trading**

50+

More than **50 research studies** have been published on RavenPack data

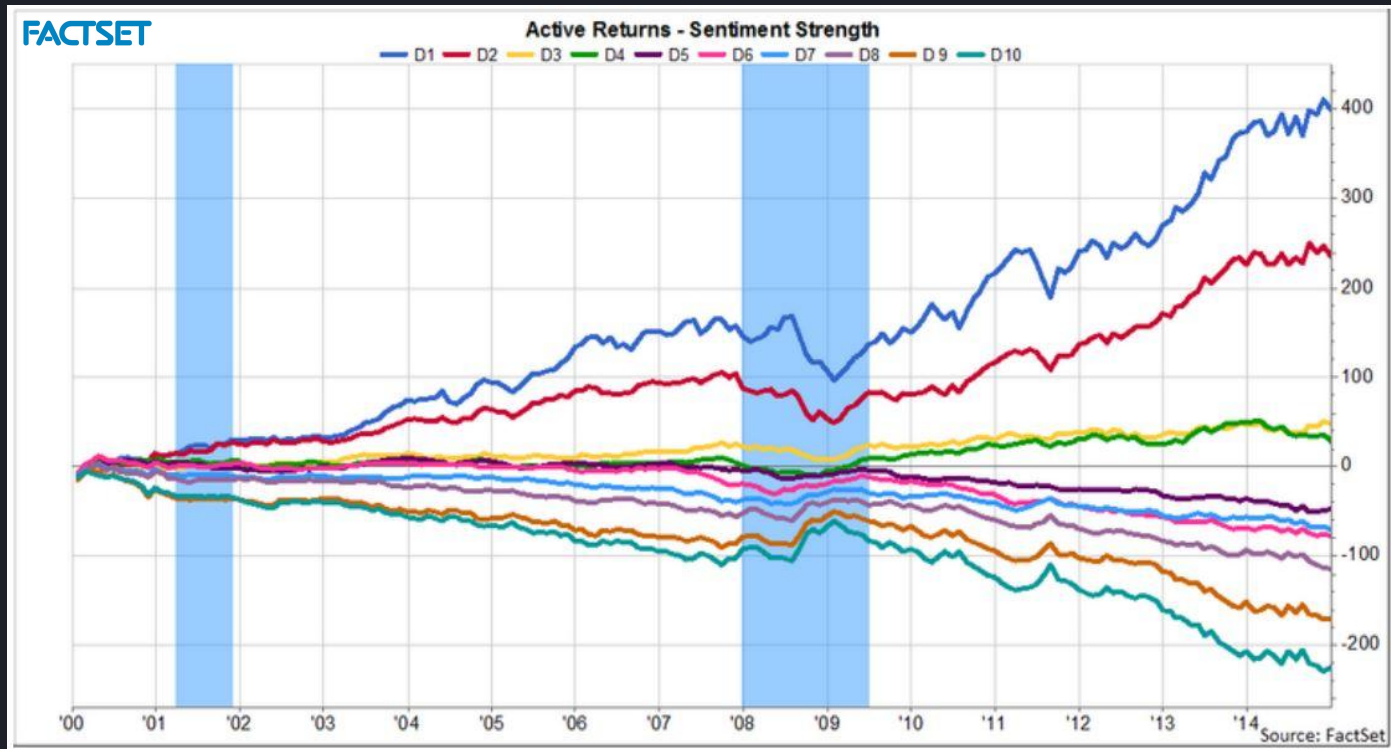
Use Cases



 RavenPack

Positive versus Negative Sentiment Stocks

Positive sentiment stocks generally outperforms negative sentiment stocks. In this case looking at monthly investment horizons for the Russell 2000.



RavenPack Analytics show **strong performance!**

*RPA delivers attractive risk-return profiles (1-day holding) - taking average sentiment across **highly novel and highly relevant events***

Statistics	U.S. Large/Mid-Cap	U.S. Small-Cap	EU Large/Mid-Cap	EU Small-Cap
Annualized Return	10.9%	30.1%	14.8%	36.1%
Annualized Vol.	4.5%	6.1%	5.7%	9.1%
Information Ratio	2.42	4.92	2.60	3.99
Avg. Portfolio Size	190	193	97	47
Maximum drawdown	6.4%	4.4%	5.9%	11.23%
Turnover	85.2%	89.5%	84.1%	91.9%

Factor Neutral performance: 1-day holding

Controlling for factor exposure (**MSCI USFAST**) leads to stronger performance

IR: 2.42 -> 3.76

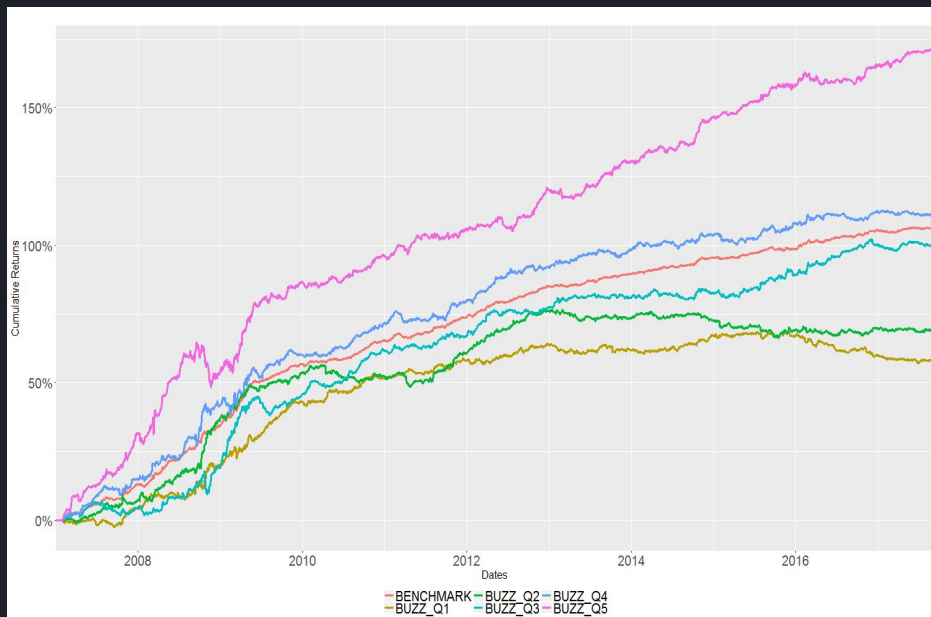
IR: 4.92 -> 6.07



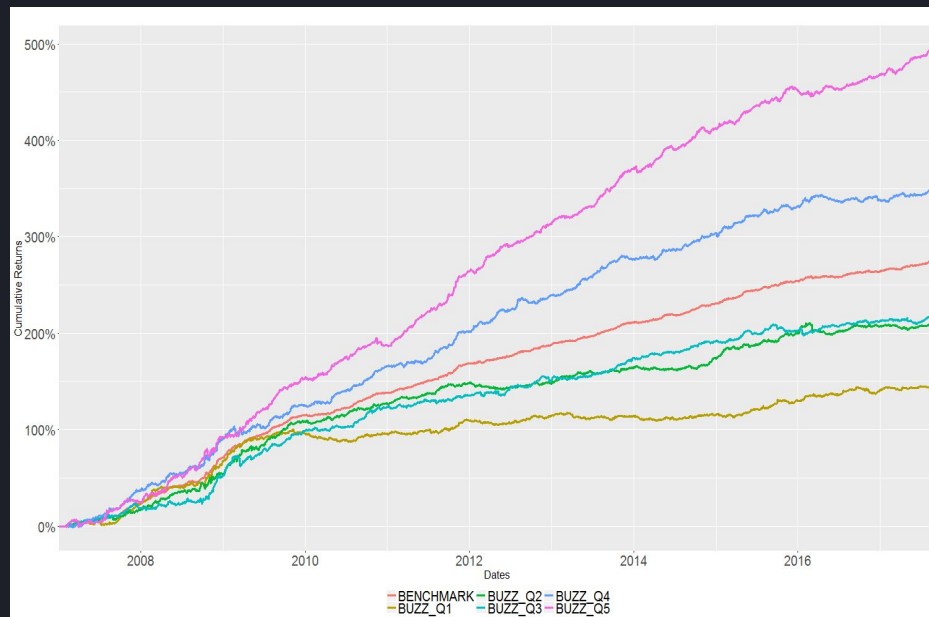
Media Buzz: high buzz companies outperform

Companies with high event buzz yield 2x greater average per-trade returns

Russell 1000, IR: 1.80 -> 2.08

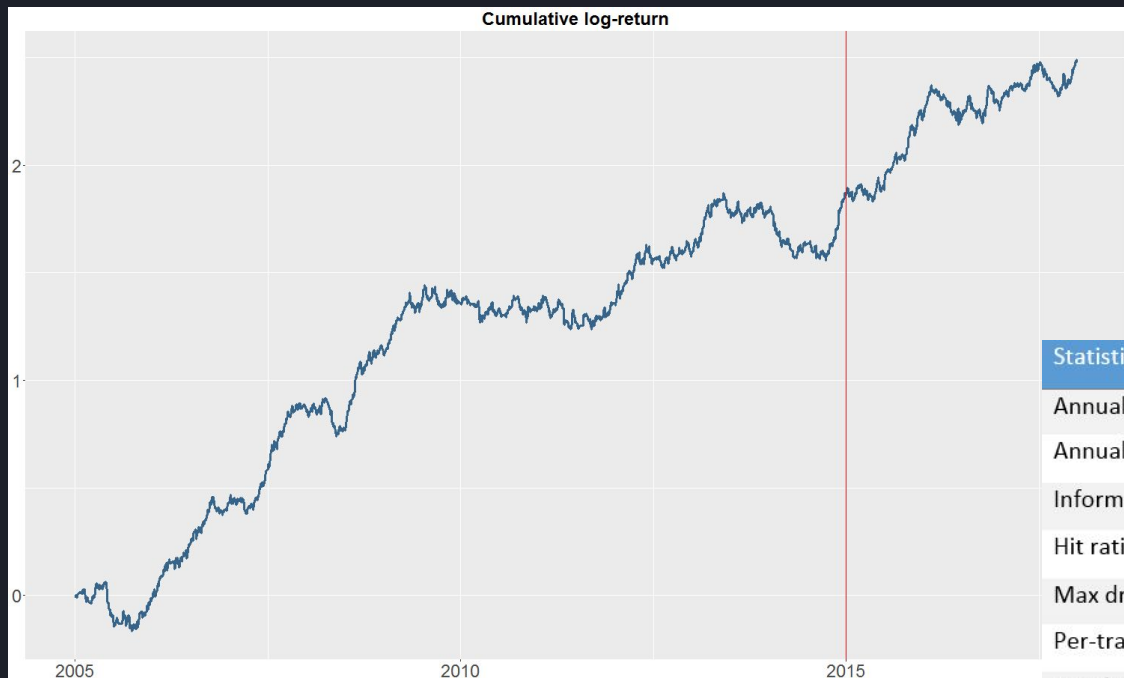


Russell 2000, IR: 2.79 -> 3.74



Energy Trading: Machine Learning Ensemble

An ensemble is created combining five Machine Learning algorithms. Further improvements are made filtering for **high** volatility driven signals.



Statistics	Ensemble	High
Annualized return	8.3%	20.5%
Annualized volatility	15.1%	17.0%
Information ratio	0.55	1.21
Hit ratio	51.1%	53.3%
Max drawdown	39.6%	18.4%
Per-trade return (basis point)	1.69	4.08
Number of trades	2740	1929

In conclusion

- The asset management industry is becoming increasingly competitive - forcing companies to identify **new avenues for alpha generation**.
- With the proliferation of social media, mobile phones/apps, and internet-of-things (IoT); more and more **(big)data** is becoming available. **Natural Language Processing** is key as lots of new content comes unstructured.
- Several **proven use cases** on how to take advantage of big and alternative data in finance and investing. The industry is waking up realizing that this is now a **“must have”!**



Questions?



RavenPack

phafez@ravenpack.com