



FIRST

Large scale information extraction and integration
infrastructure for supporting financial decision making.

My inflation pressure gauge rose at a good clip over the first half of this year on the strength of a rally in the commodity market. The commodity market recently adjusted, increased by 2.7% for the year through Aug. However, the pressure gauge has flattened out in recent months as the rally in commodities has stalled. In fact, the commodity market for '09 is being tested presently (chart).

There could be several reasons for the stall out in the upward movement. The obvious one is that with a global economy just starting from low levels, supplies are ample. But I suspect players are also being influenced by the OFTC inquiry into the trading structure of these markets as well as the recent weakness of China's market and concerns about how strongly a credit-driven business recovery may proceed.

The quick moderation of inflation thrust has helped the bond market in recent months, and it may also account for the more stable US dollar seen in the couple of weeks. Normally, the prospect of a moderation of inflation pressure as economic recovery commences would be cheered by stock investors, and we have seen that, although there could be some backlash potential as players begin to worry again about deflation, why some traders are concerned that a rising US dollar could be bad for stocks in that it would signify a flight to quality in anticipation of a failed economic recovery.

FIRST

*European research for web information extraction
and analysis for supporting financial decision making*

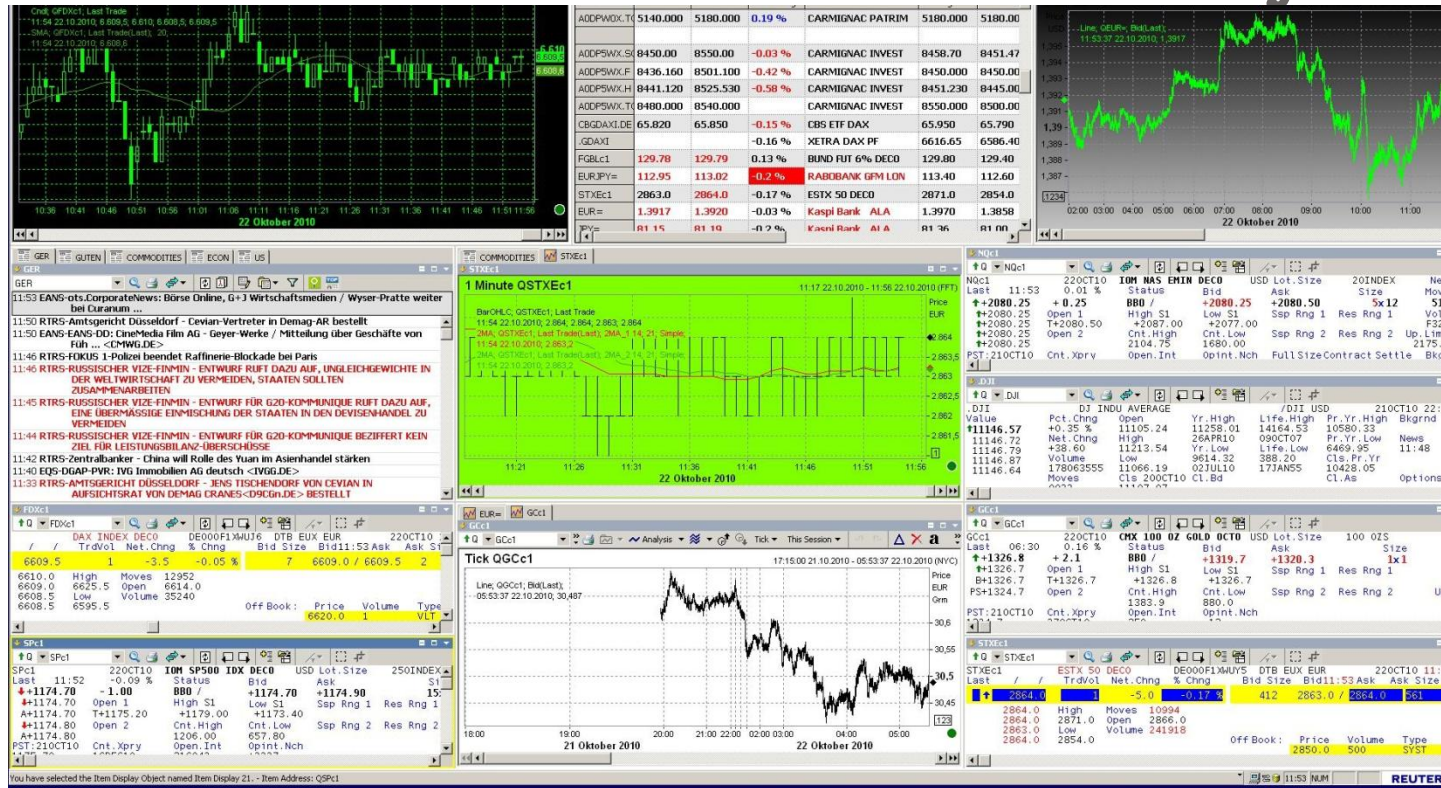
ABI Lab Forum 2012

Tomás Pariente Lobo – Atos Spain



Why FIRST? - Motivations

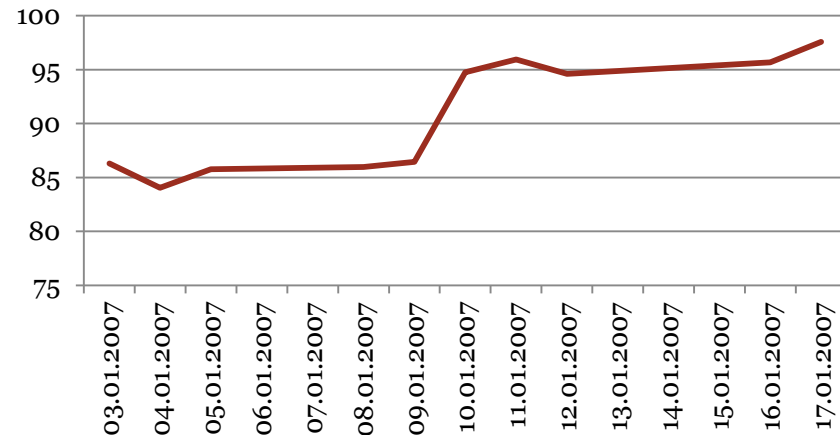
The most reliable data sources today...



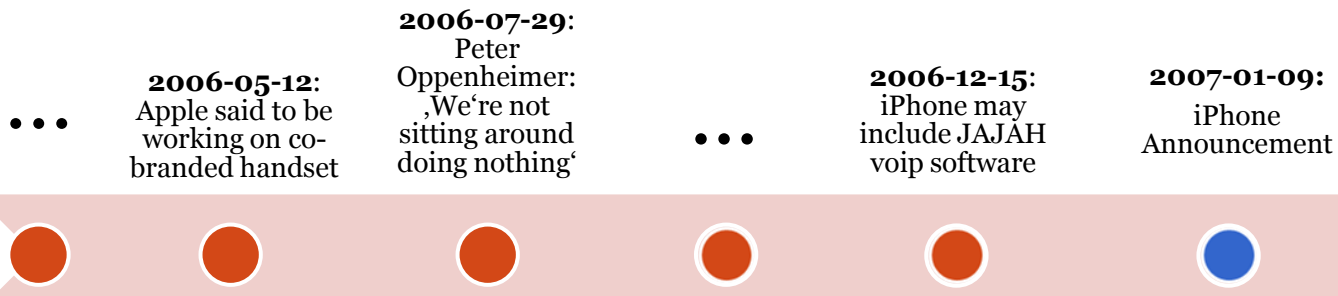
...also have their weakness!

They do not consider unstructured data, rumors, market sentiments, etc.

Example: Apple iPhone I Announcement on 2007-01-09



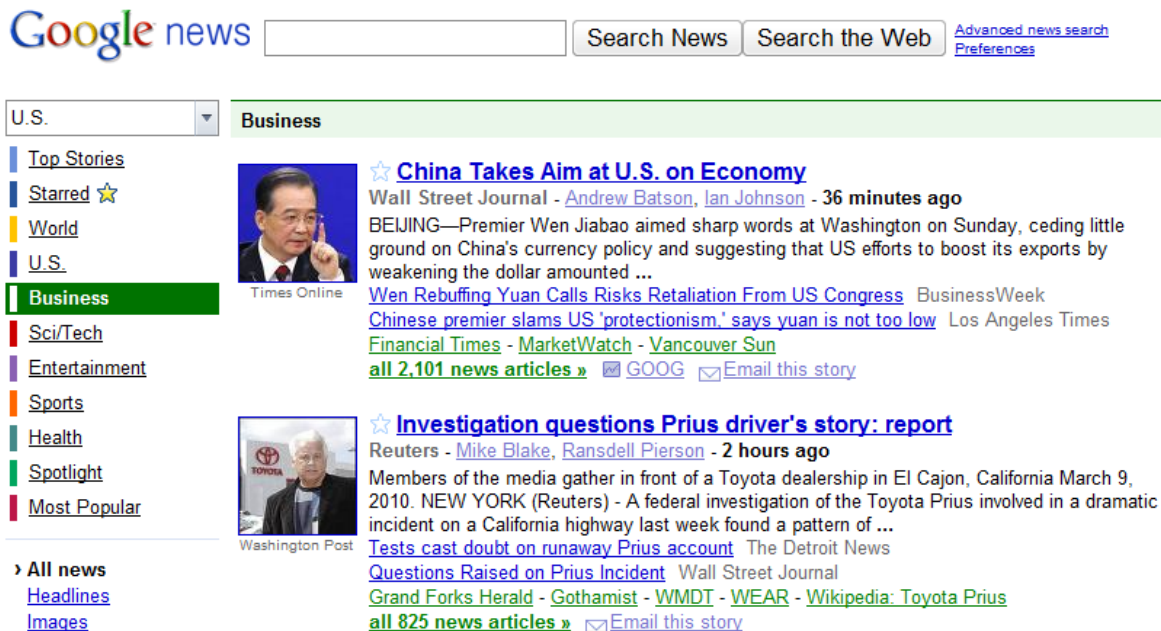
➔ **Stock prices were skyrocketing after the announcement. However, the announcement could be sensed before...**



Why FIRST? - Motivations

Example: Market surveillance via FIRST (the Google news case)

- ▶ September 2008: **Google news** announced “United Airlines bankruptcy”.
- ▶ Within 12 minutes → stock price decreased 75% → **wiped out US \$ 1bn.**



The screenshot shows the Google News interface for the 'Business' category. The left sidebar lists various topics, with 'Business' highlighted. The main content area displays two news stories. The top story is titled 'China Takes Aim at U.S. on Economy' from the Wall Street Journal, dated 36 minutes ago. The bottom story is titled 'Investigation questions Prius driver's story: report' from Reuters, dated 2 hours ago. A red arrow points from the 'Ask FIRST first' button to the top story.



- ▶ The “news” was actually **6 years old...**
- ➔ **Plausibility checking** will help in identifying hoaxes: **consistence with regulatory news and other sources.**

CalculatedRISK

FINANCE & ECONOMICS

[About](#) | [Advertising](#) | [Email CR](#) | [Graph Galleries](#) | [Tanta](#) | [Weekly Schedule](#) | [Sponsored Offers](#)

[Subscribe via email](#) | [Twitter](#) | [RSS Feed](#)

SUNDAY, JANUARY 15, 2012

Europe: Greek Debt talks will resume next week

by CalculatedRisk on 1/15/2012 09:43:00 AM

From the WSJ: [Greece to Resume Debt Talks](#)

Greece will resume talks with its private-sector creditors next week on a massive debt restructuring plan ... "Our counterparts from the Institute for International Finance will return on Wednesday and our goal is to have a general outline agreed before the next euro-group meeting on Jan. 23," Mr. Venizelos said in a speech late Saturday.

Under such a timeline, Greece would then proceed with a formal debt offer during the week of Feb. 6-10, Mr. Venizelos said, with the final debt exchange expected to be completed by the end of February.

This will be (another) critical week for Europe. The goal is to have a preliminary deal before the next

Google Custom Search

Search

Last 10 Posts

- Jan 15 at 9:43 AM **Europe: Greek Debt talks will resume next week**
- Jan 14 at 7:36 PM **Unofficial Problem Bank List and Quarterly Transition Matrix**
- Jan 14 at 4:03 PM **Schedule for Week of Jan 15th**
- Jan 14 at 8:15 AM **Summary for Week ending January 13th**
- Jan 14 at 0:19 AM **Record Home Sales in Las Vegas in 2011**
- Jan 13 at 7:43 PM **Mortgage Settlement Update**
- Jan 13 at 4:54 PM **Europe Update**

Why FIRST? – Motivations

A growing universe of unstructured data

CalculatedRISK
FINANCE & ECONOMICS

About | Advertising | E

SUNDAY, JANUA

Europe: C

by CalculatedRisk

Home Tech Entertainment Wall Street Markets Strategy Sports Lifestyle Politics Europe Data Advertising Your Money Latest Your New

Hive Tape Blackboard Contributors Charts Authors Documents Jobs Misc Video

Login With Facebook | Login With Twitter | Login | Reg

SEARCH

f Like <30k

t

From the WSJ: C **Charts**

Greece will
restructuri
return on V
euro-group
Under such
week of Fe
completed
This will be (anot



SEND US A TIP!

Get BUSINESS INSIDER Emails & Alerts

Your Email Address SIGN-UP

Learn More »

Your Money

NASDAQ Composite	2,711	+3.26	(+0.12%)
S&P 500	1,289	-5.73	(-0.443%)
NYSE Composite	7,632	-49.23	(-0.641%)

The Engage-O-Meter

Why FIRST? – Motivations

A growing universe of unstructured data

CalculatedRISK
FINANCE & ECONOMICS

About | Advertising | E

SUNDAY, JANUA

Europe: C

by CalculatedRisk | Hive | Tape | Bla

From the WSJ: C

Greece will
restructuri
return on V
euro-group

Under such
week of Fe
completed

This will be (anot

Home | Tech | Entertainment | Wall Street | Markets | Strategy | Sports | Lifestyle | Politics | Europe | Data | Advertising | Your Money | Latest | Your New

Seeking Alpha^α
Read. Decide. Invest.

Home | Portfolio | Market Currents | Long & Short Ideas | Investing for Income | ETFs & Portfolio Strategy | Macro View | Earnings | Sign In / Join Now

chart of the i

Trade deficit, in millions

\$30,000

\$20,000

\$10,000

TOP NEWS

- Shutdown of Nigerian oil output averted
- Merkel aims to hasten eurozone measures
- Obama looks to reward "insourcing"
- S&P unleashes wave of EU downgrades
- Greek debt restructuring talks collapse
- iPad 3 production underway

MARKET CURRENTS

Login With Facebook | Login With Twitter | Login | Reg

MAKE THE NEW SEEKING ALPHA HOME PAGE WORK FOR YOU!

Name the stocks you care about

ANALYSIS FOR SUNDAY, JAN 15, 10:30AM ET

Is China Now Number One?
by Scott Sumner

Investments To Cope With The 'Great Contraction'
by March Ahead

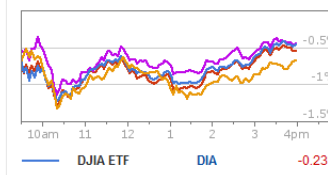
It's Time To Take Profits On These 2012 Winners
by David Alton Clark

Important Takeaways From JPMorgan's Earnings Report

Regarded Solutions' 'Buy On Dips' Dividend Strategy

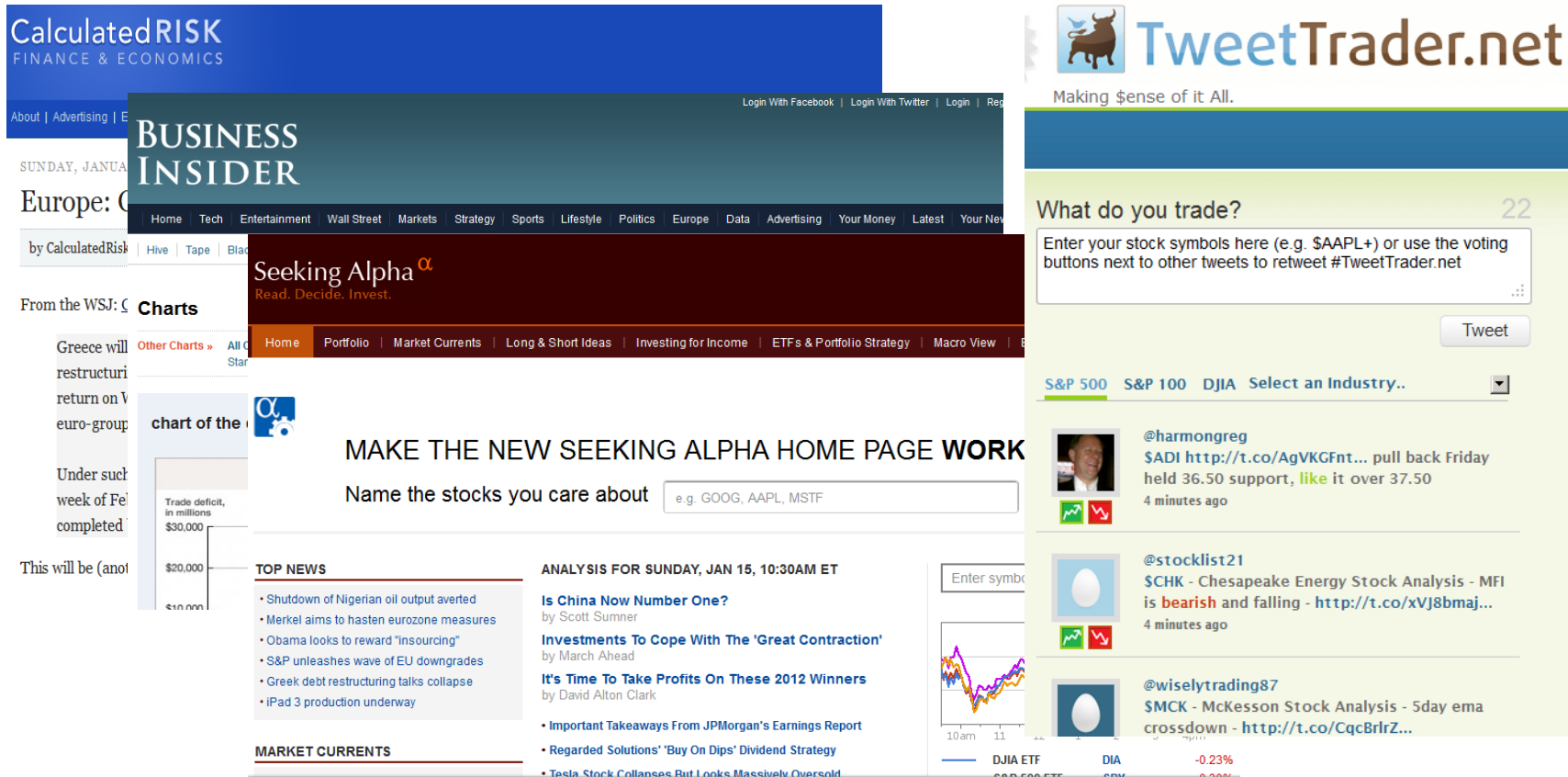
Tesla Stock Collapses But Looks Massively Oversold

Enter symbol, author, keyword



Why FIRST? – Motivations

A growing universe of unstructured data



The collage illustrates the integration of structured and unstructured data in financial analysis. It features three main components:

- CalculatedRISK FINANCE & ECONOMICS**: A blue header with navigation links (About | Advertising | E) and a search bar.
- BUSINESS INSIDER**: A dark blue header with a navigation menu (Home | Tech | Entertainment | Wall Street | Markets | Strategy | Sports | Lifestyle | Politics | Europe | Data | Advertising | Your Money | Latest | Your News) and a search bar.
- Seeking Alpha**: A dark red header with a navigation menu (Home | Portfolio | Market Currents | Long & Short Ideas | Investing for Income | ETFs & Portfolio Strategy | Macro View | B) and a search bar.

Below the headers, the collage shows various financial data and social media interactions:

- Charts**: A section titled "Greece will restructure return on V euro-group" and "Under such week of Fe completed". It includes a line chart showing "Trade deficit, in millions" from \$10,000 to \$30,000.
- TOP NEWS**: A list of headlines including "Shutdown of Nigerian oil output averted", "Merkel aims to hasten eurozone measures", "Obama looks to reward 'insourcing'", "S&P unleashes wave of EU downgrades", "Greek debt restructuring talks collapse", and "iPad 3 production underway".
- MARKET CURRENTS**: A section titled "ANALYSIS FOR SUNDAY, JAN 15, 10:30AM ET" with headlines like "Is China Now Number One?", "Investments To Cope With The 'Great Contraction'", and "It's Time To Take Profits On These 2012 Winners".
- TweetTrader.net**: A social media interface with the tagline "Making sense of it All." and a section titled "What do you trade?". It displays tweets from users like @harmongreg, @stocklist21, and @wiselytrading87, along with a "Tweet" button and a "22" count.

Why FIRST? – Motivations

A growing universe of unstructured data

CalculatedRISK
FINANCE & ECONOMICS

About | Advertising | E

SUNDAY, JANUA

BUSINESS INSIDER

Home Tech Entertainment Wall Street Markets Strategy Sports Lifestyle Politics Europe Data Advertising Your Money Latest Your New

by CalculatedRisk | Hive | Tape | Bla

From the WSJ: **Charts**

Greece will restructuri return on V euro-group

Under such week of Fe completed

This will be (anot

Other Charts » All Star

chart of the

Trade deficit, in millions

\$30,000

\$20,000

\$10,000

TOP NEWS

- Shutdown of Nigerian oil output averted
- Merkel aims to hasten eurozone measures
- Obama looks to reward "insourcing"
- S&P unleashes wave of EU downgrades
- Greek debt restructuring talks collapse
- iPad 3 production underway

MARKET CURRENTS

ANALYSIS FOR SUNDAY, JAN 15, 10:30AM ET

Is China Now Number One?
by Scott Sumner

Investments To Cope With The 'Great Contraction'
by March Ahead

It's Time To Take Profits
by David Alton Clark

- Important Takeaways From
- Regarded Solutions' 'Buy On
- Tesla Stock Collapses But L

Seeking Alpha
Read. Decide. Invest.

Home Portfolio Market Currents Long & Short Ideas Investing for Income ETFs & Portfolio Strategy Macro View

MAKE THE NEW SEEKING ALPHA HOME PAGE WORK

Name the stocks you care about

e.g. GOOG, AAPL, MSTF

Enter symbol

TweetTrader.net

Making sense of it All.

What do you trade? 22

Enter your stock symbols here (e.g. \$AAPL+) or use the voting buttons next to other tweets to retweet #TweetTrader.net

Tweet

S&P 500 S&P 100 DJIA Select an Industry..

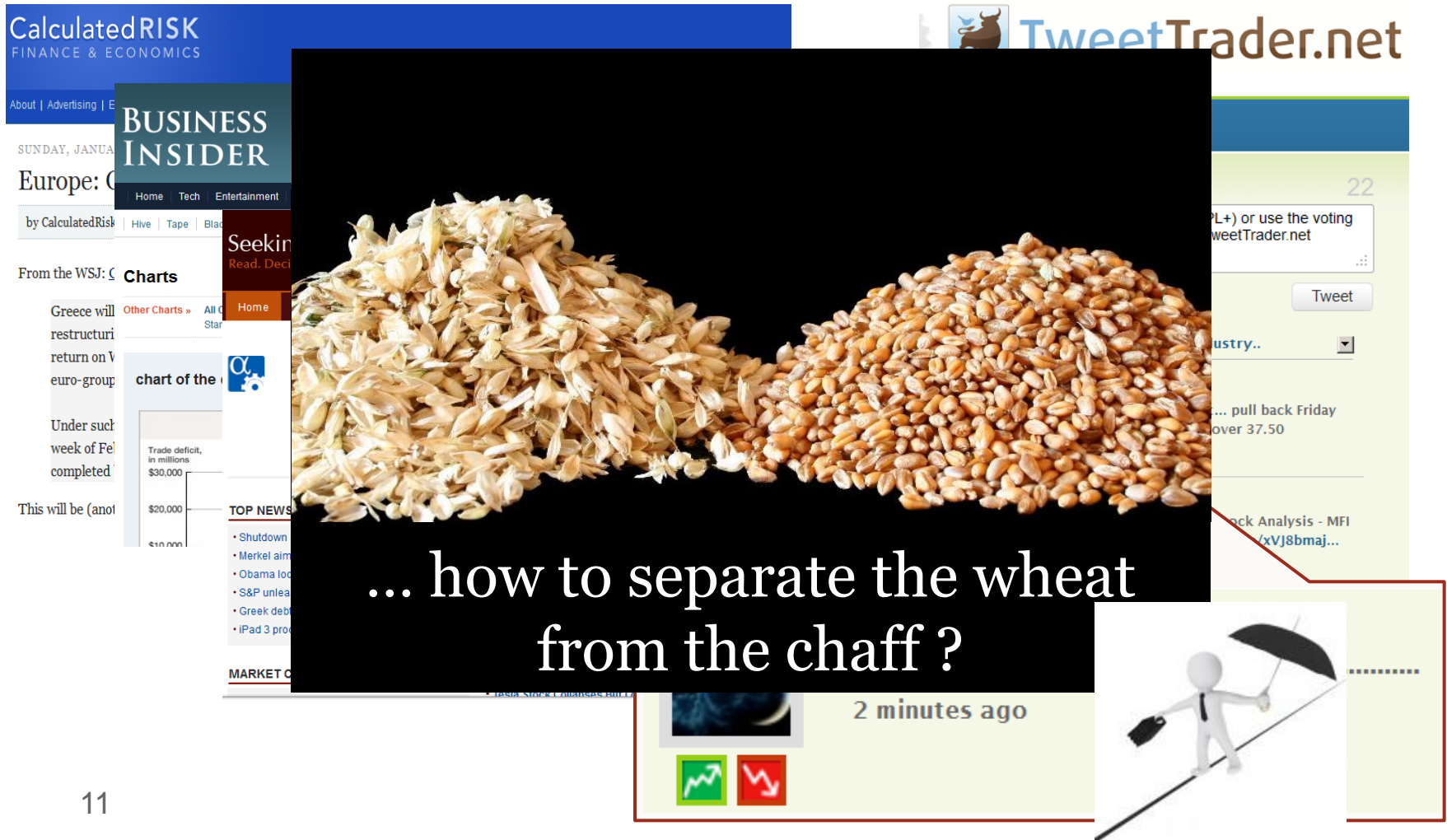
@harmongreg
\$ADI http://t.co/AgVKGfnt... pull back Friday held 3 support, like it over 37.50
4 minutes

@stockli
\$CHK - CH is bearish... /xVJ8bmaj...
4 minutes ago

@DEMO2516
\$IBM , going higher from here
2 minutes ago

Why FIRST? – Motivations

A growing universe of unstructured data



The collage features several overlapping elements:

- CalculatedRISK FINANCE & ECONOMICS**: A blue header with navigation links like 'About | Advertising | E'.
- BUSINESS INSIDER**: A dark blue header with navigation links like 'Home | Tech | Entertainment'.
- Europe: C**: A snippet of an article by CalculatedRisk.
- Charts**: A section with a line chart showing 'Trade deficit, in millions' with values from \$10,000 to \$30,000.
- TOP NEWS**: A list of news items including 'Shutdown', 'Merkel aim', 'Obama loc', 'S&P unlea', 'Greek deb', and 'iPad 3 pro'.
- MARKET C**: A snippet of a market report.
- TweetTrader.net**: A social media interface showing a tweet with a 'Tweet' button and a '22' count.
- Wheat and Chaff**: A central image showing two piles of wheat, one with the husk (chaff) and one without (wheat), illustrating the concept of separating signal from noise.
- Text Overlay**: A large white text overlay on a black background reads: **... how to separate the wheat from the chaff ?**
- Illustration**: A small illustration of a person walking a tightrope with an umbrella, symbolizing the challenge of navigating unstructured data.
- Timeline**: A snippet showing '2 minutes ago' with a green up arrow and a red down arrow.





EUROPEAN
COMMISSION

European-funded research project

Project facts

*Running from October 2010 until
September 2013*

9 partners

More than 30 people

Preliminary results available

More to come...

Stay tuned (<http://project-first.eu/>)



Who is behind FIRST?

Industrial partners



**MONTE
DEI PASCHI
DI SIENA**
BANCA DAL 1472



Academic/Research



Institut "Jožef Stefan"

SMEs

Interactive Data



b-next
CAPITAL MARKETS COMPLIANCE

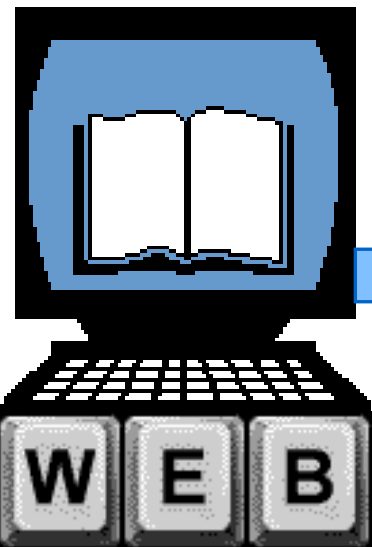
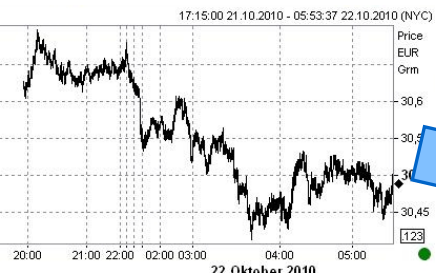




*is to **make available the relevant information**
of the entire **financial information space**
(including unreliable, unstructured, sentiment sources)
to the decision maker in near-real time
in an automated way*

Financial Resources

Structured

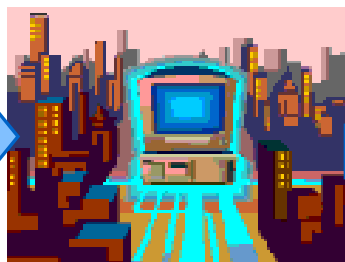


Unstructured

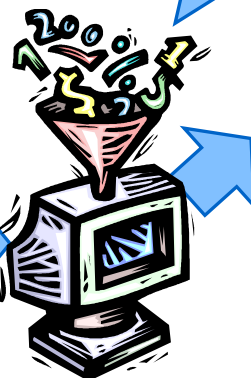
Blog, analysis, bulletin boards...
Unreliable, poor quality, noisy...

FIRST

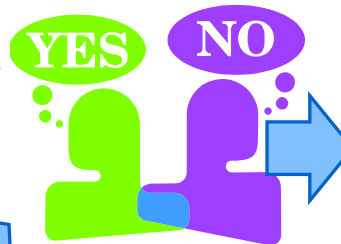
AUTOMATION



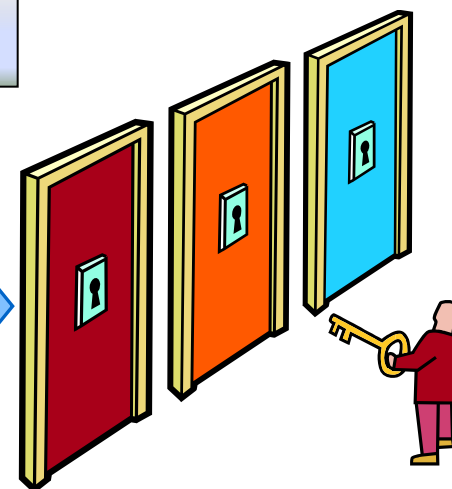
Acquisition



Processing



Analysis



Decision support



Data Acquisition pipeline: Web mining

W E B

Social
media
BLOG
NEWS



Streaming

twitter



Cleaning

CONTENT
HTML

Natural Language preprocessing
and entity extraction



ANNOTATED
CONTENT

Financial terms,
Companies,
Instruments ...

- Some numbers
 - 176 Web sites
 - 2,671 RSS sources
 - ~40,000 documents per day
 - >5,000,000 documents by end of 2011
 - And growing



Essential for future evaluation and analysis

The Analytical Pipeline: Identify, extract, classify, aggregate



Object → **Indicators** → **Positive sentiment**

Sentiment Sentences

url:
http://www.streetinsider.com/Insiders+Blog/Apple+%28AAPL%29+Could+Easily+Grow+2012+EPS+to+\$50+or+More!/6844924.html

Apple (AAPL) Could Easily Grow 2012 EPS to \$50 or More!
2011-10-10.

When Apple (Nasdaq: AAPL) reports fourth-quarter and fiscal 2011 results on October 18th, the company will show that earnings per share grew 84 percent year-over-year. For a company the size of Apple this is an astonishing feat.

However, despite this amazing growth and still limitless opportunities with the iPhone and iPad upgrades, next year analysts are only predicting growth of a modest 18 percent.

While the law of large numbers will most certainly catch up to Apple one day, they haven't yet. This morning, the company announced it had record iPhone 4S pre-orders of 1 million in the first 24 hours. Analysts said the entire weekend could have seen a total of 3 million. It wasn't even the iPhone 5!

With number like this from a product many said was a disappointment, it is easy to see that 2012 could be another Year of the Apple. Think iPad 3 and iPhone 5!

What if the company grew earnings again at 84 percent again? If this was the case you'd have a company earning \$51 per share. That's right \$51!

Slap a 20x multiple on this and you have \$1020 stock. Slap a 15x multiple on this and you have a \$765 stock. Even slap a 10x multiple and you have a \$510 stock.

Shares of Apple are up 4 percent today to \$384.57.

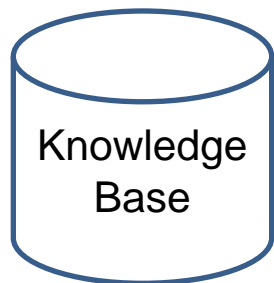
DocumentData

- Indicator
 - ☒ Indicator
- Original markups
- Sentiment
 - ☒ SentimentSentencePositive
 - SentimentPhrase
 - ☒ SentimentPhrasePositive
 - Sentiment_Object
 - ☒ SO

The Decision Support techniques: Analysis and visualization



FIRST
Acquisition &
Analytical
Pipelines



Machine
Learning
Techniques

Qualitative
Modeling

Visualization
Techniques

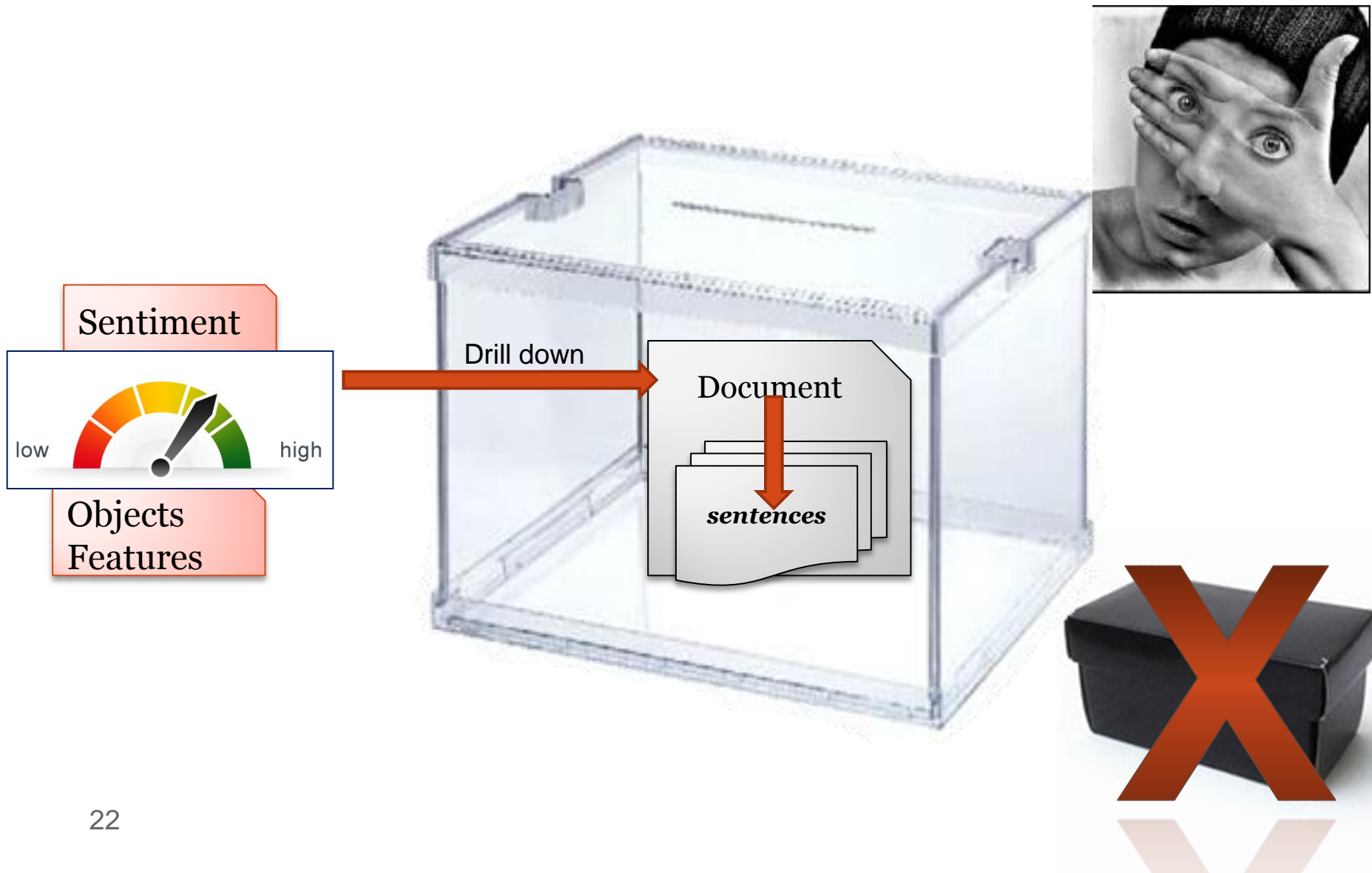
Forecasting
Models

Outputs:

Forecasts of
volatility or returns,
Alert on pump and
dump,
Reputation change
of a counterpart
Signals,
Charts,
Topic Spaces,
Topic Trends,
Reports
...



Glassbox model





The three FIRST use cases & their relevance for the industry



Market Surveillance

- **Capital markets compliance** can be automated today **using structured data**, but the automation **does not take unstructured data into account**
- **FIRST will**
 - make use of **large volumes of unstructured data** into financial compliance;
 - develop automated techniques to **better detect market abuse/insider trading**..



Reputational Risk Management

- **No off-the-shelf solutions or methodologies** for reputational risk management.
- **FIRST will**
 - provide a **sustainable tool for reputational risk monitoring**;
 - contribute to break new ground in this field of dramatically **high impact in FSI**.



Retail Brokerage

- **Today**, mainly based on **quantitative analysis and key figures**.
- **FIRST will**
 - **use unstructured data to leverage** both **information** for private investors and sophisticated **tools** for professional users.





FIRST

**Large scale information extraction and integration
infrastructure for supporting financial decision making.**

My inflation pressure gauge rose at a good clip over the first half of this year on the strength of a rally in the commodity market. In commodity market, seasonally adjusted, increased by 2.7% for the year through Aug. However, the pressure gauge has flattened out in recent months as the rally in commodities has stalled. In fact, the commodity market for '09 is being tested presently (chart).

There could be several reasons for the stall out in the uptick in commodities. The obvious one is that with the global economic recovery just starting from low levels, supplies are ample. But I suspect players are also being influenced by the OFTC inquiry into the trading structure of these markets as well as the recent weakness of China's market and concerns about how strong a credit-driven business recovery may proceed.

The quick moderation of inflation thrust has helped the bond market in recent months, and it may also account for the more stable US dollar seen in a couple of weeks. Normally, the prospect of a moderation of inflation pressure as economic recovery commences would be cheered by stock investors, and we have seen that, although there could be some backlash potential as players begin to worry again about deflation. It is why some traders are concerned that a rising bond market could be bad for stocks in that it would signify a flight to quality in anticipation of a failed economic recovery. It is a bit of a puzzle, but interesting nonetheless.

Acknowledgement

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n°257928.

THANKS