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

FIRST Project Public Presentation

Copyright © 2010, FIRST Consortium

- **Why FIRST?**
- **What is FIRST?**
- **Who is behind FIRST?**
- **FIRST Vision**
- **FIRST Innovations**
- **Expected Outcomes**
- **Impact**

Large scale inFormation extraction and Integration infRastructure for SupporTing financial decision making

The most reliable data sources today…

…also have their weakness!
They do not consider unstructured data, rumors, market sentiments, etc.
Example 1: Apple iPhone 1 Announcement on 2007-01-09
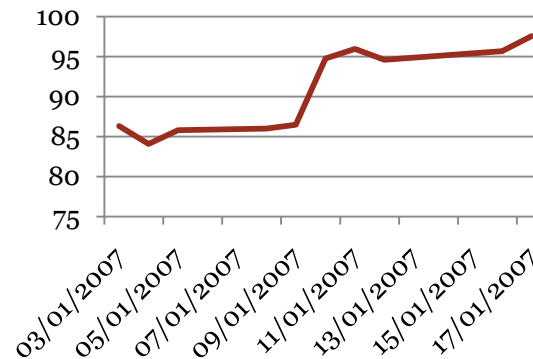
➔ **Stock prices were skyrocketing after the announcement. However, the announcement could be sensed before…**

Timeline of iPhone Rumors

...

2006-05-12: Apple said to be working on co-branded handset

2006-07-29: Peter Oppenheimer: ‚We're not sitting around doing nothing'

...

2006-12-15: iPhone may include JAJAH voip software

2007-01-09: Official iPhone 1 Announcement

2006-07-29:
Peter Oppenheimer: ‚We're not sitting around doing nothing'

2006-05-12:
Apple said to be working on co-branded handset

2006-12-15:
iPhone may include JAJAH voip software

2007-01-09:
iPhone Announcement

· · ·

· · ·

FILED UNDER *Cellphones, Portable Audio / Video*

Apple said to be working with Softbank on "co-branded" handset

By Evan Blass 🔊 *posted May 12th 2006 5:16PM*

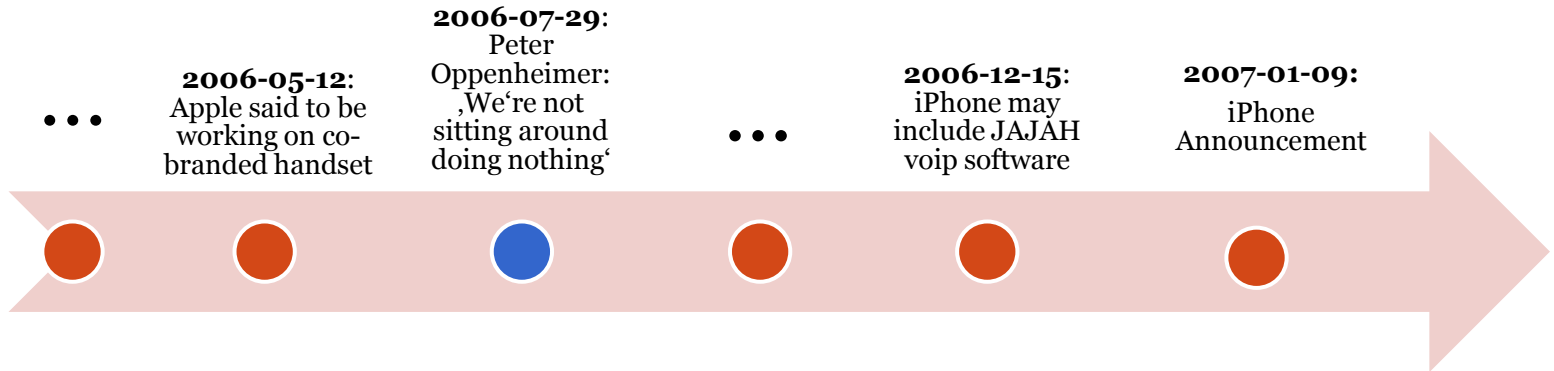Japanese news agency Nikkei is reporting that Apple and new Vodafone Japan owner Softbank are working together to develop cellphones capable of downloading and playing music from the iTunes store. Details are, of course, very slim at this point, as we don't even know whether to characterize this as a "true" iPhone, or simply a ROKR-like model featuring only software from the iPod manufacturer powering a non-Apple handset. However, Nikkei is quoting sources that say the phone, which may be released as early as this year (and which is definitely NOT pictured alongside this post), is expected to sport both Softbank and Apple branding, which may indicate that Cupertino will have a hand in developing the hardware as well. Whatever they're working on, we wanna know about it, and we won't rest until we have a ton of specs and pics for you.

6

Why FIRST? - Motivations

2006-05-12:
Apple said to be working on co-branded handset

2006-07-29:
Peter Oppenheimer: ‚We're not sitting around doing nothing'

2006-12-15:
iPhone may include JAJAH voip software

2007-01-09:
iPhone Announcement

Apple Phone In August? [Updated]

Saturday July 29, 2006 09:33 PM EST
Written by Jeff Longo

Contrary to recent reports, Engadget says that Apple's rumored "iPhone" may appear as early as August.

A reader is reporting to us that a coworker's tech-unsavvy friend, who is regularly hired by Apple to do marketing photo shoots, was recently brought on to take some shots of "the sleekest, sexiest damn phone he's ever seen."

It has been well-established that Apple has been working on an Apple-branded phone for some time, however the release date has been difficult to pinpoint. Recently, various patents (1, 2) have been uncovered regarding Apple's cell phone technology and Peter Oppenheimer made comments during Apple's Q3 2006 conference call virtually acknowledging Apple's work on the product, stating "We're not sitting around doing nothing" with regard to cell phones.
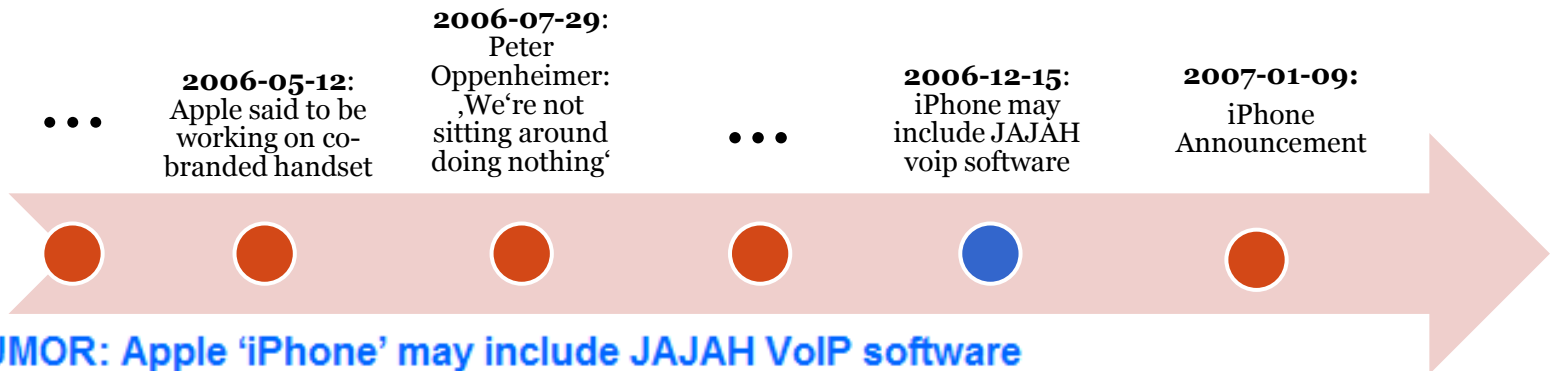
Update: Engadget points to an AppleInsider post revealing that the June 28th iPod updater package has some Phone related text strings such as "t_feature_app_PHONE_APP, kPhoneSignalStrength, clPhoneCallModel, clPhoneCallHistoryModel, prPhoneSettingsMenu".

Mac Guides: iPhone (Rumored)
Submitted by menziep
Rating (282 Positives; 67 Negatives)
[422 comments]

2006-07-29:
Peter
Oppenheimer:
‚We're not
sitting around
doing nothing'

2006-05-12:
Apple said to be
working on co-
branded handset

2006-12-15:
iPhone may
include JAJAH
voip software

2007-01-09:
iPhone
Announcement

• • •   • • •

RUMOR: Apple 'iPhone' may include JAJAH VoIP software

Friday, December 15, 2006 – 03:59 PM EDT

Daniel Mattes, co-founder of Mountain View, CA-based VoIP software developer JAJAH, has been quoted as saying that JAJAH is presently holding intense conversations with Apple over the use of JAJAH software for a mobile phone, according to a reports from the German-language website "futureZone." JAJAH has been described as a "Skype Killer" by some (one example and another).

When contacted by futureZone, JAJAH officials would not confirm or deny working with Apple, saying only, "We have nothing to announce."

According to the JAJAH website:
JAJAH is a new type of phone company... JAJAH is dedicated to bringing the world vastly improved telephony solutions at a fraction of the traditional price. Whenever possible, we want to provide basic phone for free - and we truly believe it's possible. We believe that by bringing together the best of the internet with the best of the traditional telephone industry we will be able to provide our customers with unique new solutions, solutions never before possible, at a price they'll find irresistible.

Unlike other companies that are trying to replace the phone with the computer, JAJAH believes that phones are great the way they are. Phones work. People are not interested in headsets, downloads or hotspots. They simply want to make a call. And we completely agree.

8

Example 2: Market surveillance via FIRST (The Google case)

▸ September 2008: **Google news** announced "United Airlines **bankruptcy**".

▸ Within 12 minutes → stock price decreased 75% → **wiped out US $ 1bn.**

▸ **The "news" was actually 6 years old…**

→ **Plausibility checking** will help in identifying hoaxes: **consistence with regulatory news and other sources**.

So "Why FIRST?":

- Financial actors are extraordinarily faced with the challenges of extremely large, dynamic and heterogeneous information bases.

- Need for integrating online unstructured data analysis with the current analysis on financial structured data.

- Dependence on availability of highly scalable, extremely flexible, and performing methods for semantic information extraction, information integration, and decision support.

- FIRST addresses issues of the financial and economic crisis:
  - hidden risks,
  - hidden information,
  - information asymmetries,
 - and dynamically evolving conditions.

10

What is FIRST?

- FIRST (FP7-257928) is a European Commission FP7 funded Specific Targeted Research Project (STREP)

- FIRST started on October 1st, 2010 and finishes on September 30th, 2013.

- The overall budget is roughly 4,6 million euro with a maximum EU contribution of 3,025 million euro.

> Overall Objective
> *Develop and provide **ICT infrastructure** for **Extracting, Integrating, and Exploiting** yet unused **massive amounts of unstructured data** (e.g., textual data, web information) **for financial decision making***

FIRST Vision

Vision

*The FIRST vision is to **improve** existing methods **and**, where necessary, **to provide new methods** for*

semantic information extraction, information integration, and decision support

applied to the financial domain

so that the entire financial information space

can be explored and made use of dynamically, i.e.,

at near real-time*.*

FIRST Vision

Financial Resources
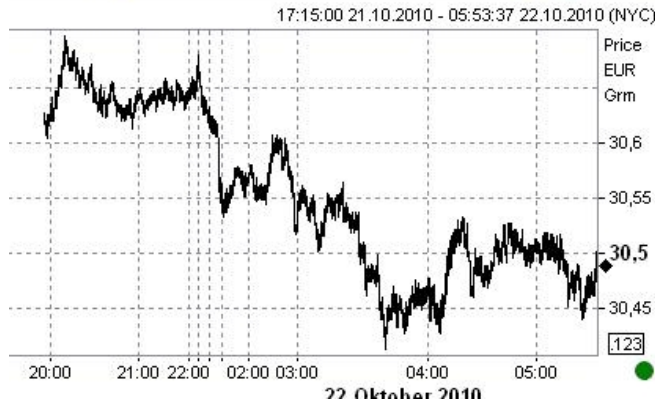
News article · Expert analysis · Blog entries · Bulletin Boards
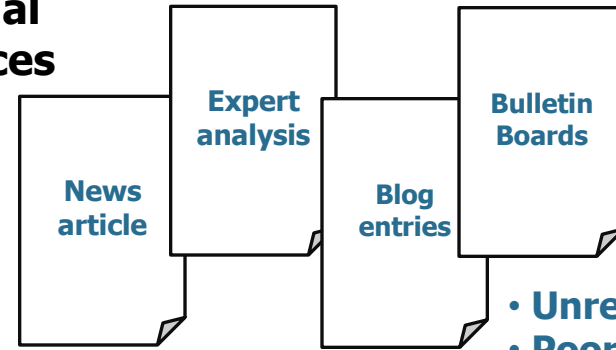
- Unreliable
- Poor quality
- Noisy

structured → **unstructured** →

Manually impossible to collect, filter and assess these types and amounts of information.
The solution comes with automation!
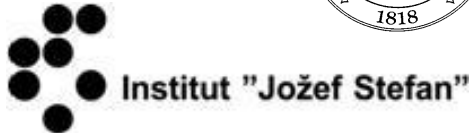
↓

FIRST

Exploits and integrates financial information that is **dynamic**, **extremely large**, and **heterogeneous** by using methods that will address noise and uncertainty in order to scale the information and support in the financial decision making process
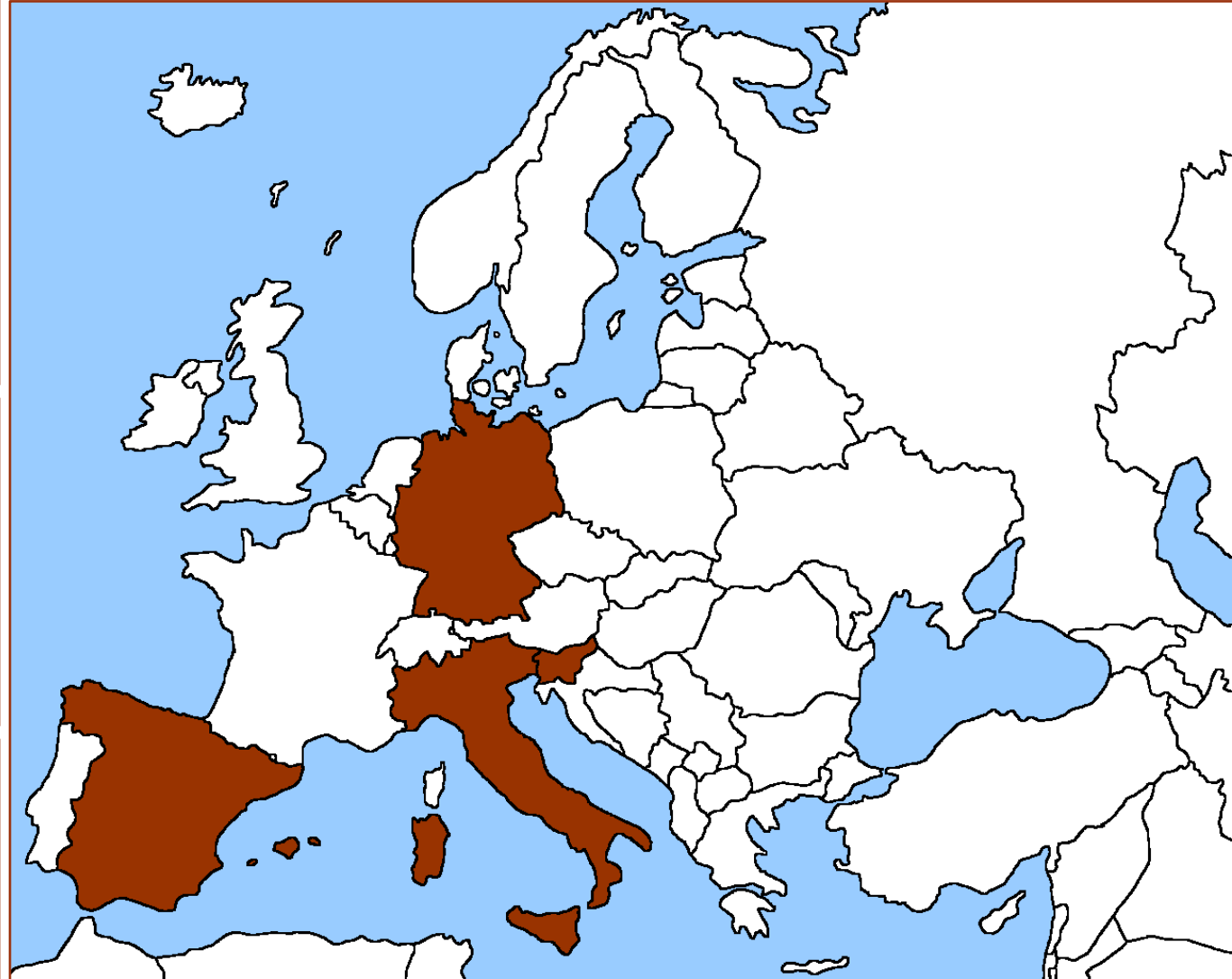
13

Industrial partners

Atos Origin

MONTE DEI PASCHI DI SIENA
BANCA DAL 1472

Boerse Stuttgart

Academic/Research

GOETHE UNIVERSITÄT
FRANKFURT AM MAIN

UNIVERSITÄT HOHENHEIM
1818

Institut "Jožef Stefan"

SMEs

b-next
ENGINEERING GMBH

Interactive Data

Who is behind FIRST?

Information Extraction

Sentiment Analysis

Decision Support

The composition of the FIRST Consortium guarantees that **scientific, technological and development objectives of the project will be met**

The consortium includes:

3 research partners providing complementary know-how and expertise in knowledge modeling,, semantic information extraction, and decision support infrastructures:

- **Institut Jozef Stefan, Slovenia**
- **University of Hohenheim, Germany**
- **Goethe University Frankfurt, Germany**

Semantic Technologies

Knowledge Engineering

4 distinct use case providers (market surveillance, reputational risk management, and online retail banking and brokerage) which exploit on end-user prototypes extending the financial market information system. The results will be evaluated by real end-users, customers (of FIRST partners), and members of the FIRST Advisory Board.

- **b-next engineering GmbH, Germany**
- **Banca Monte dei Paschi di Siena, Italy**
- **Interactive Data Management Solutions, Germany**
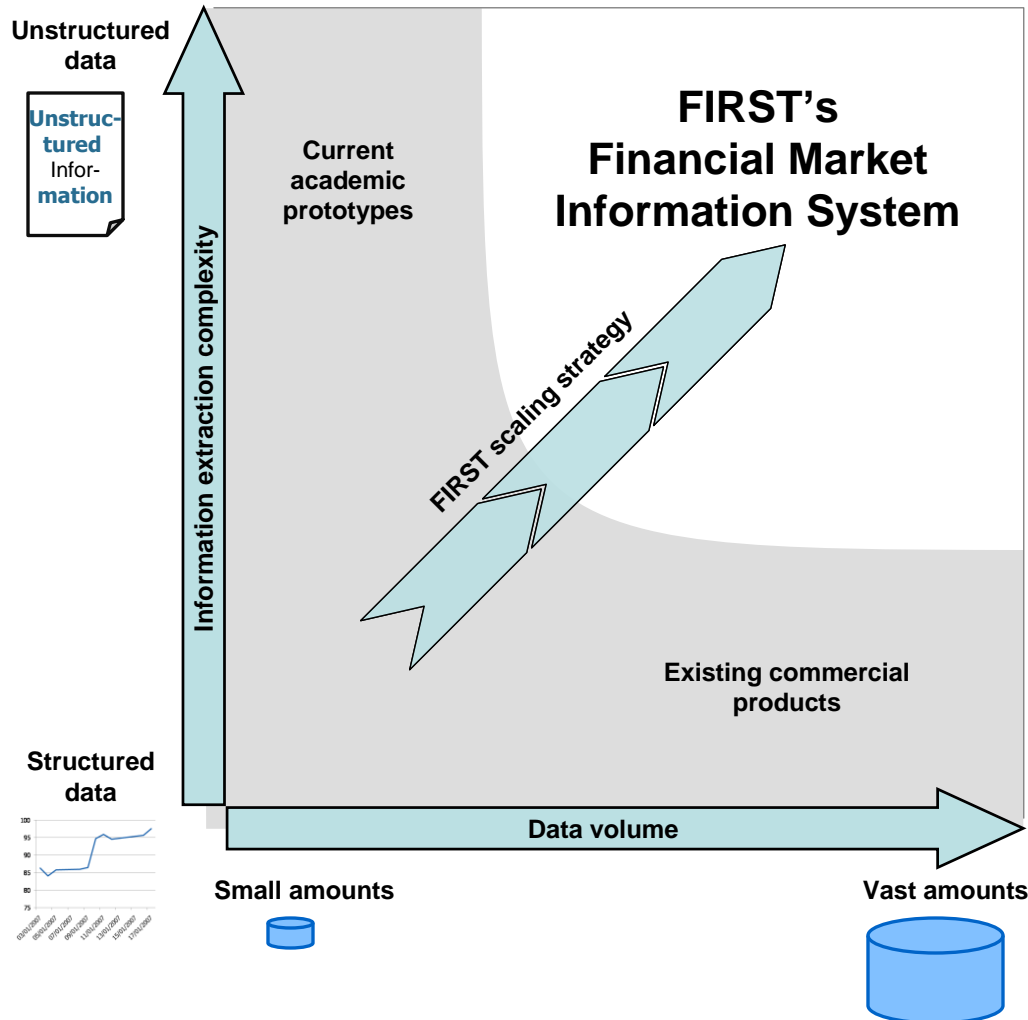- **Boerse Stuttgart Holding GmbH, Germany**

1 industrial consulting partner, project coordinator and playing the roles of software integrator, financial market portal open source solution provider, and researching on semantic technologies:

- **Atos Origin S.A.E., Spain**

Industrial Exploitation

Financial Data and Software Providers

Financial Institutions

1. **Structured → Unstructured Data**
   - Address noise and uncertainty

2. **Small → Vast** amounts of data

3. **Offline → Online streams (near real time)**
   - Ontology infrastructure
   - Machine learning
   - Sentiment extraction
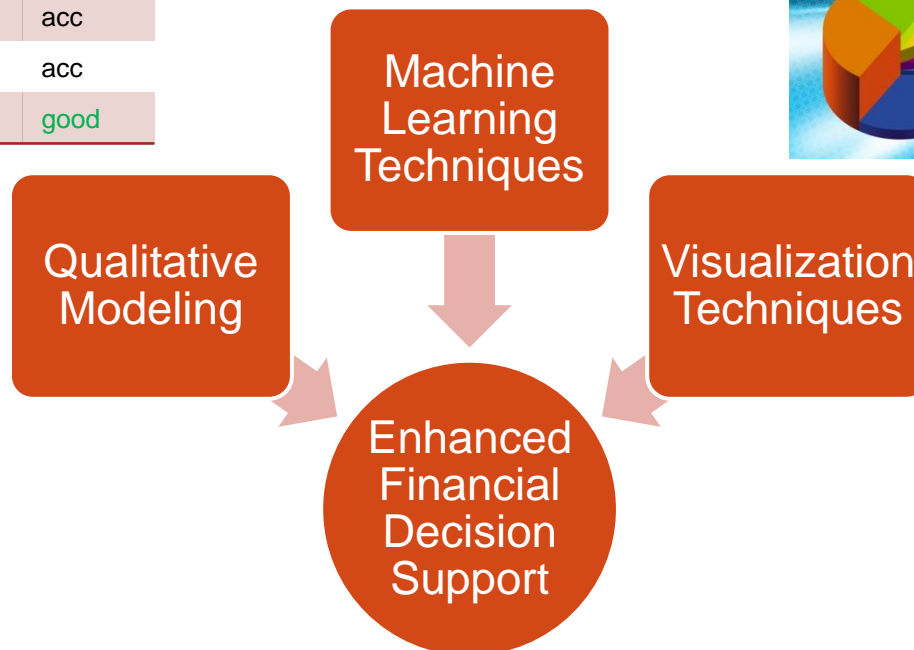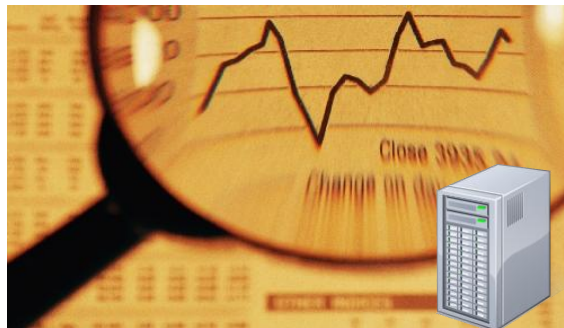   - Qualitative modeling
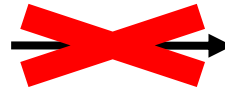   - Visualization

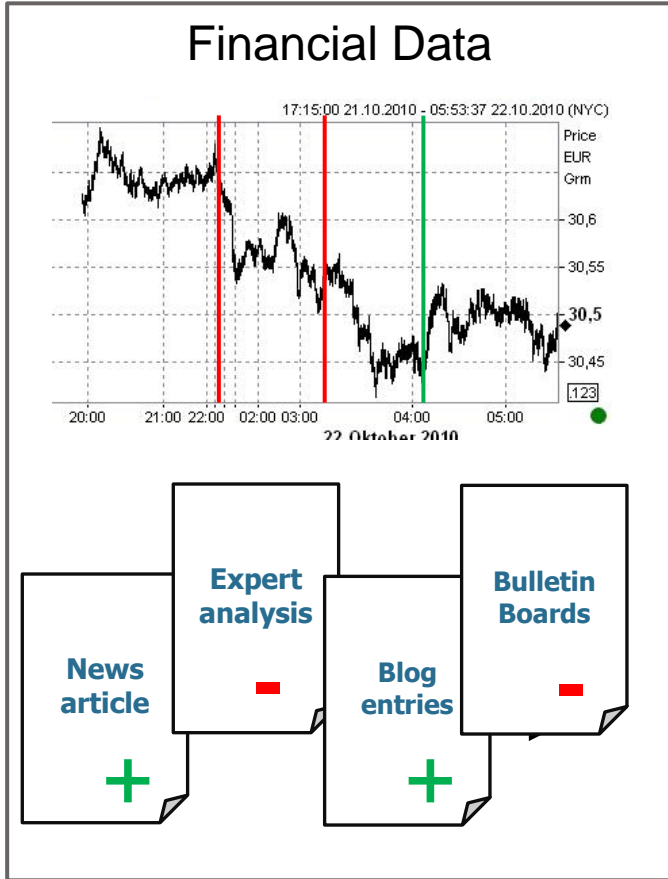4. **Financial decision support**
   - Based on high-level semantic features
   - Glass-box models

- Deep **semantic** sentiment analysis on **multiple levels** (starting with words and sentences) for higher accuracy
- **Well grounded**, shaded sentiment being more close to reality with clear interpretation (possibility, probability)
- Incorporating **new** web information **sources** enabling access to information before-the-fact
- **Glass-box** model for traceability
- Addressing **noise and uncertainty**
- **Active learning** for sentiment vocabulary acquisition
- **Real-time** sentiment analysis
- Use sentiment-related high-level features in **event-detection** and prediction models

FIRST will follow a multi-method approach to enhance financial decision making

Price	Vola	P&L	Assess
high	low	good	unacc
med	high	good	unacc
med	Med	good	acc
low	Low	acc	acc
low	Low	good	good

Qualitative Modeling

Machine Learning Techniques

Visualization Techniques

Enhanced Financial Decision Support

Financial Data

17:15:00 21.10.2010 - 05:53:37 22.10.2010 (NYC)

Price EUR Grm

30,6
30,55
30,5
30,45

20:00 21:00 22:00 02:00 03:00 04:00 05:00

22. Oktober 2010

News article

Expert analysis

Blog entries

Bulletin Boards

high-level features

Advanced financial decision making

Data visualization

Decision support models

Expected Outcomes
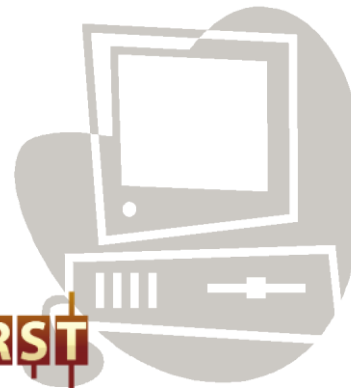
- Real-time and scalable financial unstructured data acquisition, extraction, analysis, integration, and visualization pipeline

- Ontology infrastructure, active learning

- Well grounded, deep semantic sentiment analysis

- Advanced financial decision support (qualitative glass box) models based on high-level features

- Advanced real-time visualization techniques

- Open source and commercial version of a financial market information system

FIRST Use Cases

FIRST validates its innovations on three extremely challenging and complementary use cases:

1. Market surveillance,
2. Risk management,
3. Online retail brokerage and investment management

New information-based service models.

Increase market transparency.

| Early access to new information. | Indentify unfolding potential risks. | Reduce information overflow. |

Scaling of information extraction.

Address unstructuredness, noise, and uncertainty.

How to focus on the relevant while making financial decisions in a world of information overkill?

Ask FIRST first

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

Thanks