Predictive Text Analytics

Seth Grimes
Alta Plana Corporation
301-270-0795 -- *http://altaplana.com* -- @SethGrimes

Predictive Analytics World
October 20, 2009

Context Counts

Hedging With Options

Example: You expect to receive 100,000 CAD in 3 months and want to lock in a minimum

rate at which to sell CAD against USD. You buy a CAD put:
Current Spot Rate USD/CAD: 1.3700
Strike Price: 1.3761
Maturity: 3 months
Style: European
Premium: 1.22%

This option gives you the right, but not the obligation, to sell CAD at 1.3761 at maturity.

Your cost for this option is USD $886.56

Scenarios at Maturity with an option hedge:
CAD appreciates: USD/CAD =1.2500

You choose not to exercise your option because you can sell your USD/CAD at the prevailing market rate. Net of the premium you receive is $79,113.44.

CAD depreciates: USD/CAD = 1.4900

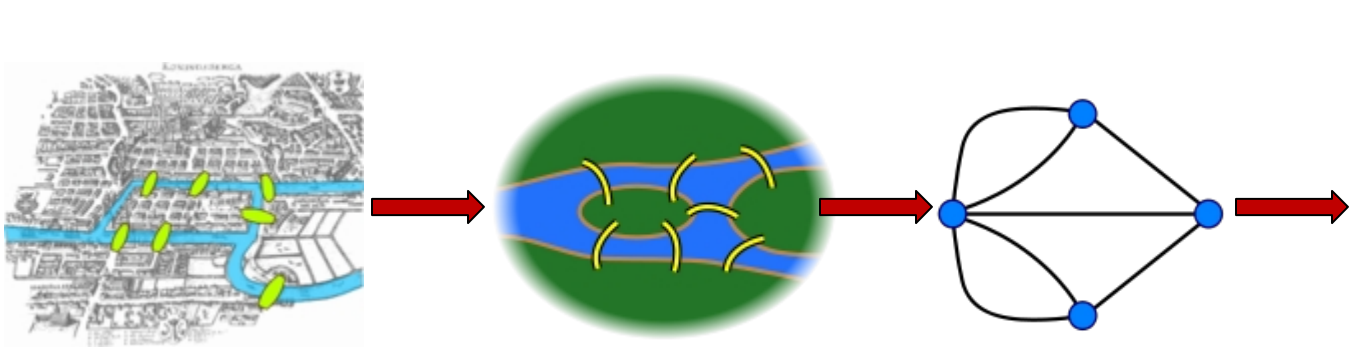
You choose to exercise your option and sell your CAD at 1.3761, receiving $72,669.14 versus the prevailing market rate where you would only receive $67,114.09. Net of the premium you receive is $71,782.58.

Eastern Bank

What is Analytics?

$$x(t) = t$$
$$y(t) = \tfrac{1}{2}\, a\, (e^{t/a} + e^{-t/a})$$
$$= a\cosh(t/a)$$

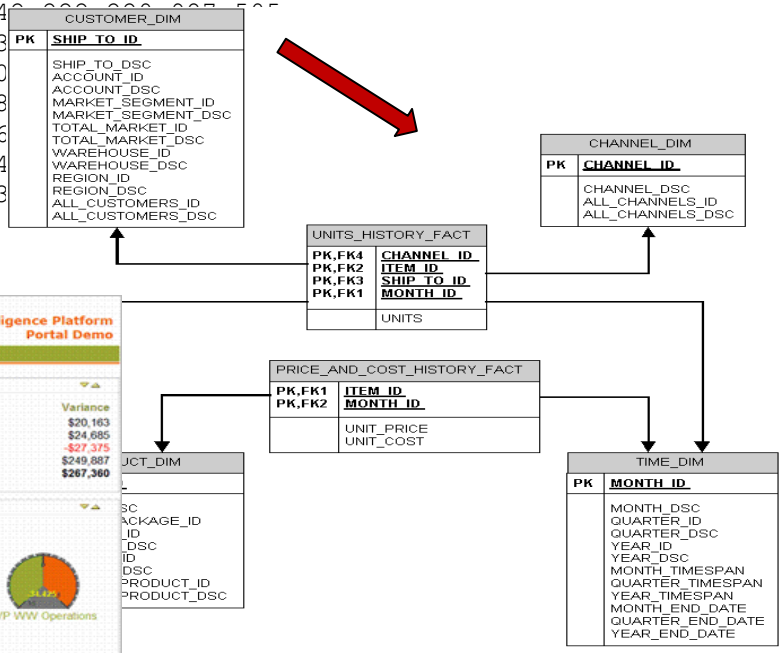
http://www.tropicalisland.de/NYC_New_York_Brooklyn_Bridge_from_World_Trade_Center_b.jpg



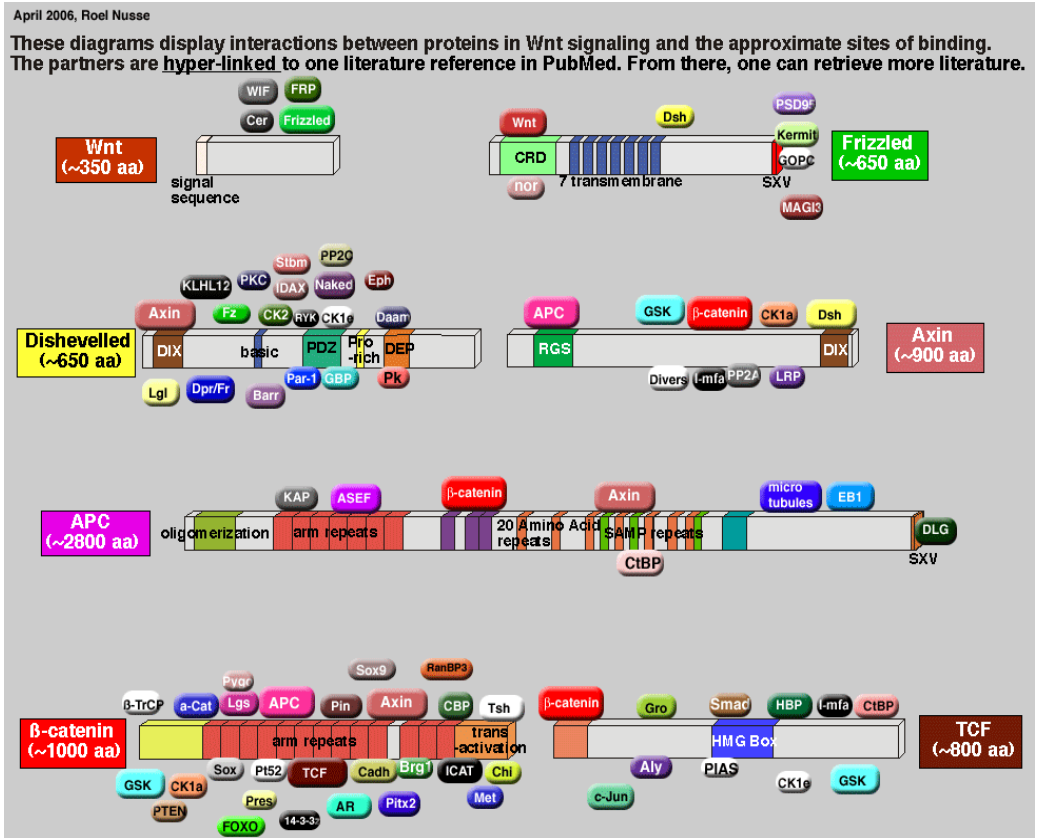
http://en.wikipedia.org/wiki/Seven_Bridges_of_K%C3%B6nigsberg

What is Analytics?

```
"SUMLEV","STATE","COUNTY","STNAME","CTYNAME","YEAR","POPESTIMATE",
50,19,1,"Iowa","Adair County",1,8243,4036,4207,446,225,221,994,509
50,19,1,"Iowa","Adair County",2,8243,4036,4207,446,225,221,994,509
50,19,1,"Iowa","Adair County",3,8212,4020,4192,44
50,19,1,"Iowa","Adair County",4,8095,3967,4128,43
50,19,1,"Iowa","Adair County",5,8003,3924,4079,40
50,19,1,"Iowa","Adair County",6,7961,3892,4069,38
50,19,1,"Iowa","Adair County",7,7875,3855,4020,36
50,19,1,"Iowa","Adair County",8,7795,3817,3978,34
50,19,1,"Iowa","Adair County",9,7714,3777,3937,33
```



The "Unstructured Data" Challenge



www.stanford.edu/%7ernusse/wntwindow.html

Axin and Frat1 interact with dvl and GSK, bridging Dvl to GSK in Wnt-mediated regulation of LEF-1.

Wnt proteins transduce their signals through dishevelled (Dvl) proteins to inhibit glycogen synthase kinase 3beta (GSK), leading to the accumulation of cytosolic beta-catenin and activation of TCF/LEF-1 transcription factors. To understand the mechanism by which Dvl acts through GSK to regulate LEF-1, we investigated the roles of Axin and Frat1 in Wnt-mediated activation of LEF-1 in mammalian cells. We found that Dvl interacts with Axin and with Frat1, both of which interact with GSK. Similarly, the Frat1 homolog GBP binds Xenopus Dishevelled in an interaction that requires GSK. We also found that Dvl, Axin and GSK can form a ternary complex bridged by Axin, and that Frat1 can be recruited into this complex probably by Dvl. The observation that the Dvl-binding domain of either Frat1 or Axin was able to inhibit Wnt-1-induced LEF-1 activation suggests that the interactions between Dvl and Axin and between Dvl and Frat may be important for this signaling pathway. Furthermore, Wnt-1 appeared to promote the disintegration of the Frat1-Dvl-GSK-Axin complex, resulting in the dissociation of GSK from Axin. Thus, formation of the quaternary complex may be an important step in Wnt signaling, by which Dvl recruits Frat1, leading to Frat1-mediated dissociation of GSK from Axin.

www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed&cmd=Retrieve&list_uids=10428961&dopt=Abstract

Modelling Text

- Metadata
  - E.g., title, author, date
- Statistics
  - Typically via vector space methods
  - E.g., term frequency, cooccurrence, proximity
- Linguistics
  - Lexicons, gazetteers, phrase books
  - Word morphology, parts of speech, syntactic rules
  - Larger-scale structure including discourse
- Machine learning

Significant words in descending order of frequency (common words omitted).

46	*nerve*	12	*body*	6	*disturbance*	4	*accumulate*
40	*chemical*	12	*effects*	6	*related*	4	*balance*
28	*system*	12	*electrical*	5	*control*	4	*block*
22	*communication*	12	*mental*	5	*diagram*	4	*disorders*
19	*adrenalin*	12	*messengers*	5	*fibers*	4	*end*
18	*cell*	10	*signals*	5	*gland*	4	*excitation*
18	*synapse*	10	*stimulation*	5	*mechanisms*	4	*health*
16	*impulses*	8	*action*	5	*mediators*	4	*human*
16	*inhibition*	8	*ganglion*	5	*organism*	4	*outgoing*
15	*brain*	7	*animal*	5	*produce*	4	*reaching*
15	*transmission*	7	*blood*	5	*regulate*	4	*recording*
13	*acetylcholine*	7	*drugs*	5	*serotonin*	4	*release*
13	*experiment*	7	*normal*			4	*supply*
13	*substances*					4	*tranquilizing*

Total word occurrences in the document:	2326
Different words in document:	
Total of different words	741
Less different common words	170
Different non-common words	571
Ratio of all word occurrences to different non-common words	~4:1
Non-common words having a frequency of occurrence of 5 and over:	
Total occurrences	478
Different words	39

163

IBM JOURNAL • APRIL 1958

"Statistical information derived from word frequency and distribution is used by the machine to compute a relative measure of significance, first for individual words and then for sentences. Sentences scoring highest in significance are extracted and printed out to become the auto-abstract."

H.P. Luhn, *The Automatic Creation of Literature Abstracts*, IBM Journal, 1958.

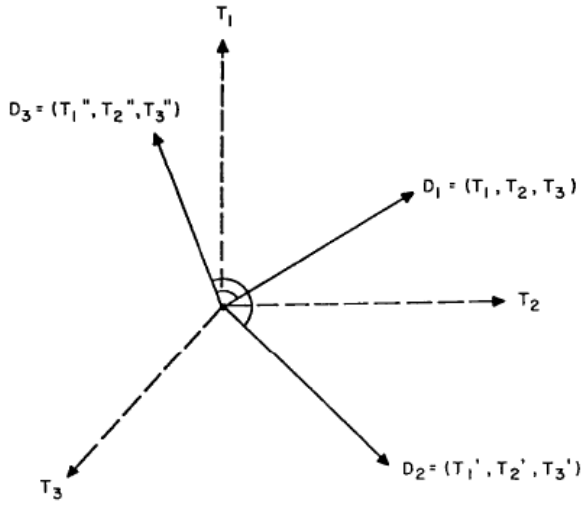
Text Modelling

Document content can be considered an unordered "bag of words."

Particular documents are points in a high-dimensional vector space.

Salton, Wong & Yang, "A Vector Space Model for Automatic Indexing," November 1975.

Fig. 1. Vector representation of document space.



Text Modelling

We might construct a term-document matrix...

D1 = "I like databases"

D2 = "I hate hate databases"

	I	like	hate	databases
D1	1	1	0	1
D2	1	0	2	1

http://en.wikipedia.org/wiki/Term-document_matrix

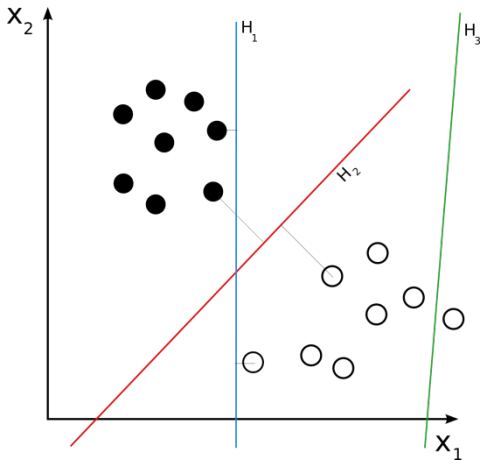
and use a weighting such as TF-IDF (term frequency–inverse document frequency)…

in computing the cosine of the angle between weighted doc-vectors to determine similarity.

Text Modelling

Analytical methods make text tractable.

- Latent semantic indexing utilizing singular value decomposition for term reduction / feature selection.
  - Creates a new, reduced concept space.
  - Takes care of synonymy, polysemy, stemming, etc.
- Classification technologies / methods:
  - Naive Bayes.
  - Support Vector Machine.
  - K-nearest neighbor.



Text Modelling

In the form of *query-document similarity*, this is Information Retrieval 101.

See, for instance, Salton & Buckley, "Term-Weighting Approaches in Automatic Text Retrieval," 1988.

A useful basic tech paper: Russ Albright, SAS, "Taming Text with the SVD," 2004.

Given the complexity of human language, statistical models may fall short.

Semantics

"This rather unsophisticated argument on 'significance' avoids such linguistic implications as grammar and syntax... No attention is paid to the logical and semantic relationships the author has established."

-- Hans Peter Luhn, 1958

New York Times,
September 8, 1957

Anaphora / coreference: **"They"**

SCIENCE IN REVIEW

Chemistry Is Employed in a Search for New Methods to Conquer Mental Illness

By ROBERT K. PLUMB

By coincidence this week-end in New York City marks the end of the annual meeting of the American Psychological Association and the begining of the annual meeting of the American Chemical Society.

Psychologists and chemists have never had so much in common as they now have in new studies of the chemical basis for human behavior. Exciting new finds in this field were also discussed last week in Iowa City, Iowa, at the annual meeting of the American Physiological Society and at Zurich, Switzerland, at the Second International Congress for Psychiatry.

Two major recent developments have called the attention of chemists, physiologists, physicists and other scientists to mental diseases: It has been found that extremely minute quantities of chemicals can induce hallucinations and bizarre psychic disturbances in normal people, and mood-altering drugs (tranquilizers, for instance) have made long-institutionalized people amenable to therapy.

Money to finance research on the physical factors in mental illness is being made available. Progress has been achieved toward the understanding of the chemistry of the brain. New goals are in sight.

At the psychiatrists meeting in Zurich last week, four New York City physicians urged their colleagues to broaden their concept of "mental disease," and to probe more deeply into the chemistry and metabolism of the human body for answers to mental disorders and their prevention.

Blood May Tell

Dr. Felix Marti-Ibanez and three brothers, Dr. Mortimer D. Sackler, Dr. Raymond R. Sackler and Dr. Arthur M. Sackler cited evidence that the blood chemistry of victims of schizophrenia is different from that of normal people. Perhaps multiple biological factors are responsible for this chemical change, they suggested.

Mental disease is a "developmental process" and long duration of a disorder may result in "permanent alteration of anatomy and physiology," they said. They urged that trials of new drugs which affect the brain should be concentrated on complex studies of the mechanism of action of the drugs. The variety of substances capable of producing profound mental effects is a new armory of weapons for use in investigating biological mechanisms underlying mental disease, they said.

The sources of behavioral disturbance are many and they may come from external as well as internal forces, the four reported. This concept has already proven practical, for instances, when it enabled psychiatrists to predict that the administration of ACTH and cortisone could produce psychosis.

"It led some years ago to the development of a blood test which was 80 per cent accurate in the identification of schizophrenic patients," they said. "It permitted us on physiologic grounds to deny that the psychoneuroses and the psychoses were lesser and greater degrees of the same disease process, and, in fact, to affirm that they represented opposite and even mutually exclusive directions of physiologic disturbances," they said.

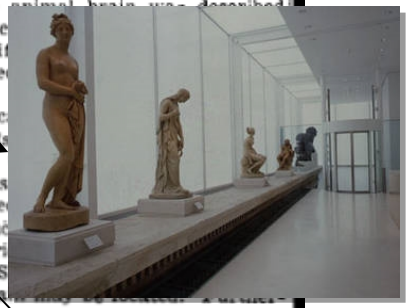
Chemicals now available should be used not only to bring relief to the mentally sick but also to uncover the biological mechanisms of the disease processes themselves. "Only then will the metabolic era mature and bring to fruition man's long hoped for salvation from the ravages of mental disease," they reported.

Chemistry of the Brain

At the psychologist's meeting here, a technique for tracing electrical activity in specific portions of the animal brain was described by rese
of Cali
reporte
in cat
electric
animals
which
In this
reporte
sequenc
its vari
ing. S
the bra
more, the electrical pathways so traced out can be blocked temporarily by the use of chemicals. This poses new possibilities for studying brain c
and sic
the Ca
sized.

The r
try hav
peutic
courage
for me
that kn
ary fiel
last we
Washin

This
the Na
Health
formati
Literat
field an
bers ca
technic
People
vited to
or othe
letters
have in
Informa
Silver



Semantics

Why do we need linguistics?

The Dow *fell* 46.58, or 0.42 percent, to 11,002.14. The Standard & Poor's 500 index fell 1.44, or 0.11 percent, to 1,263.85, and the Nasdaq composite *gained* 6.84, or 0.32 percent, to 2,162.78.

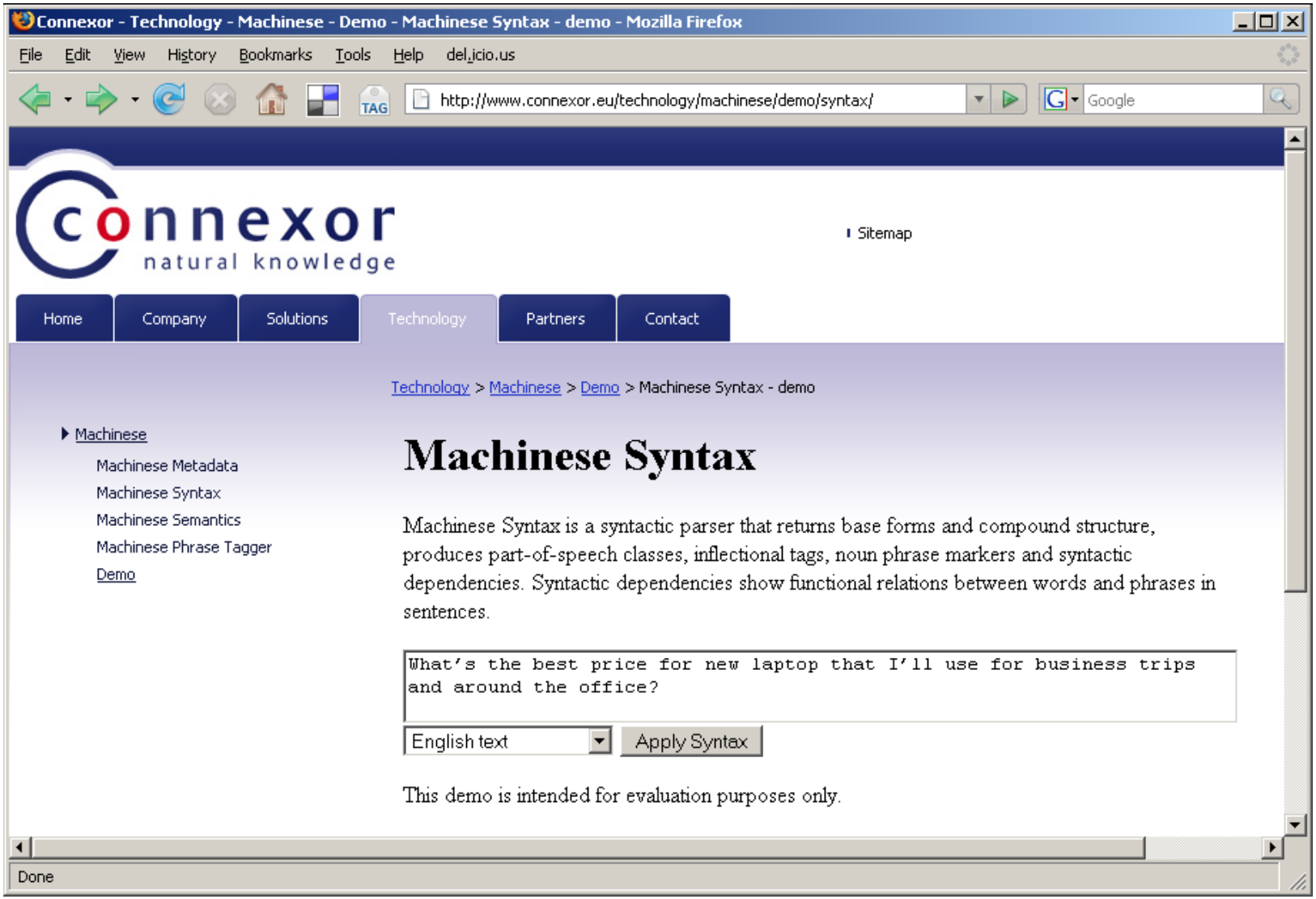
The Dow *gained* 46.58, or 0.42 percent, to 11,002.14. The Standard & Poor's 500 index fell 1.44, or 0.11 percent, to 1,263.85, and the Nasdaq composite *fell* 6.84, or 0.32 percent, to 2,162.78.

(Example: Luca Scagliarini, Expert System.)

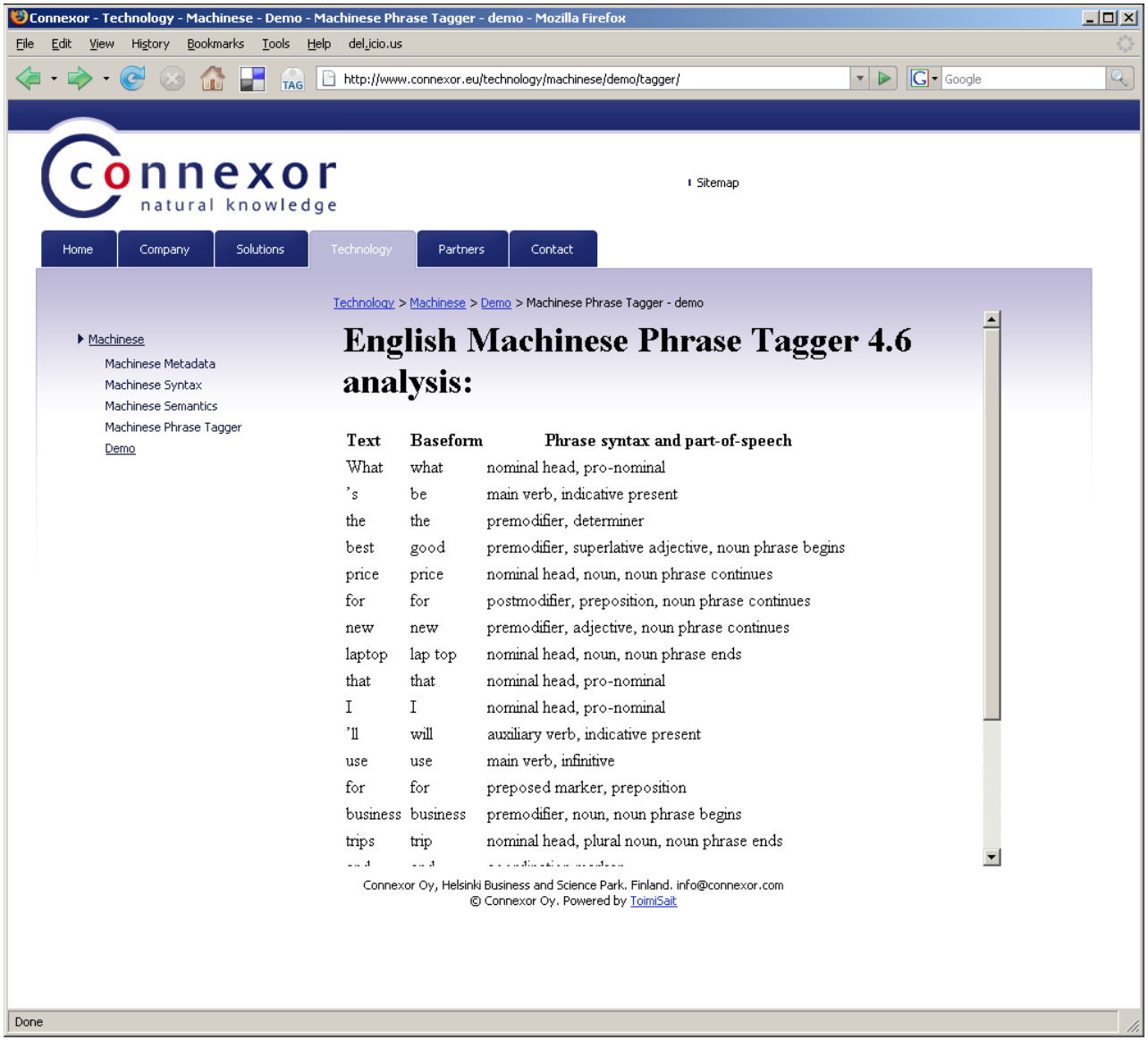
John pushed *Max*. *He* fell.

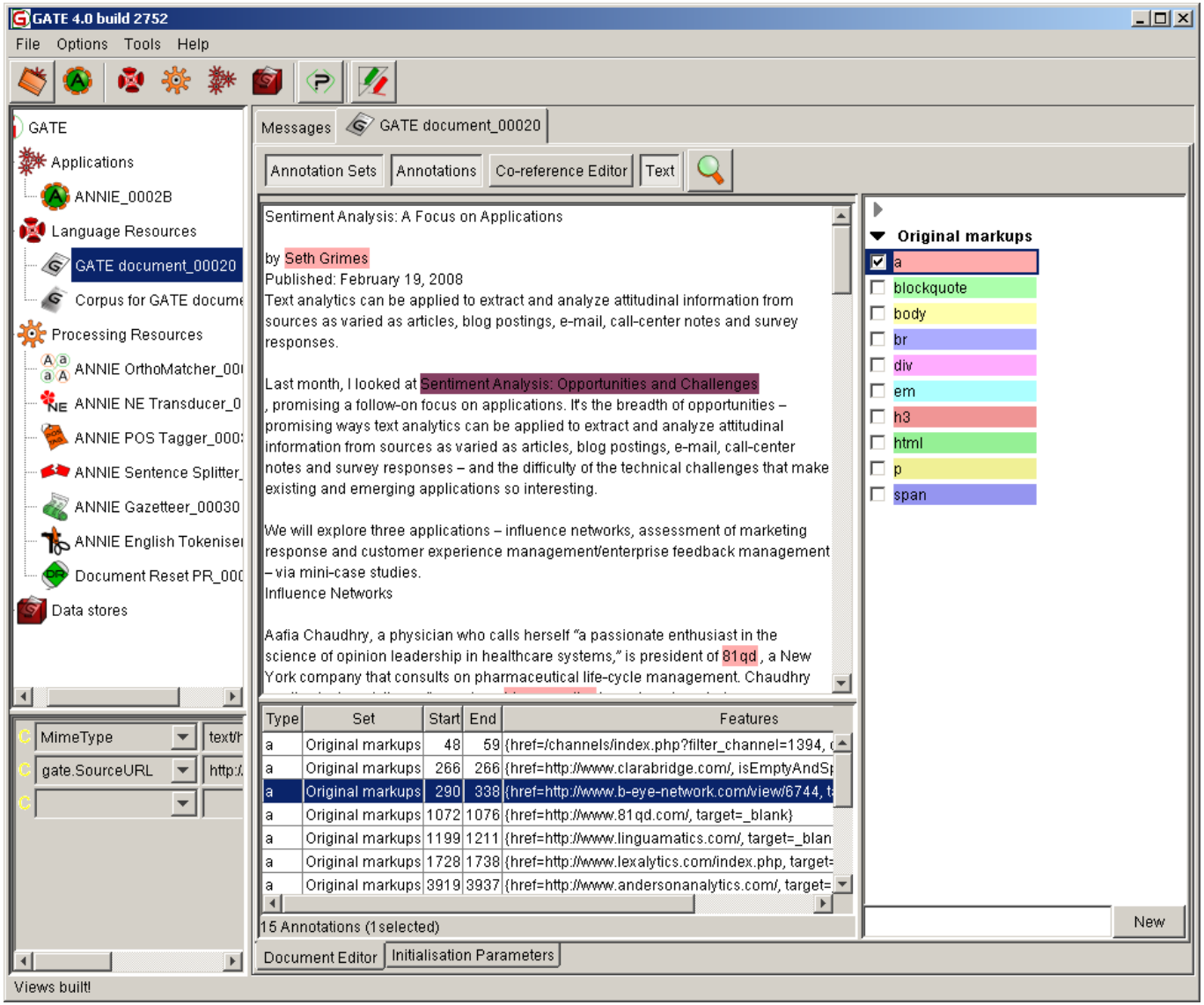
John pushed Max. *He* laughed.

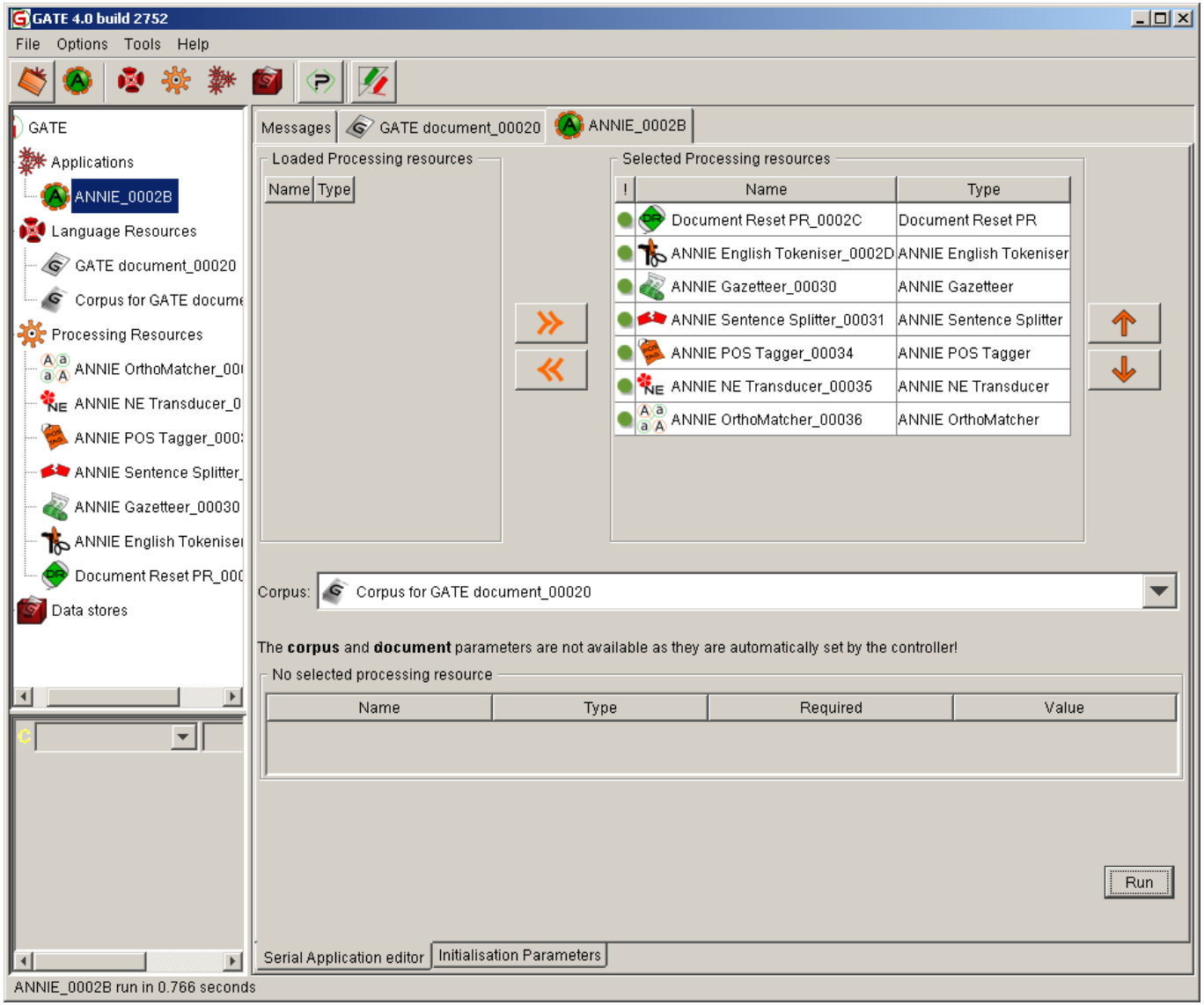
(Example: Laure Vieu and Patrick Saint-Dizier.)

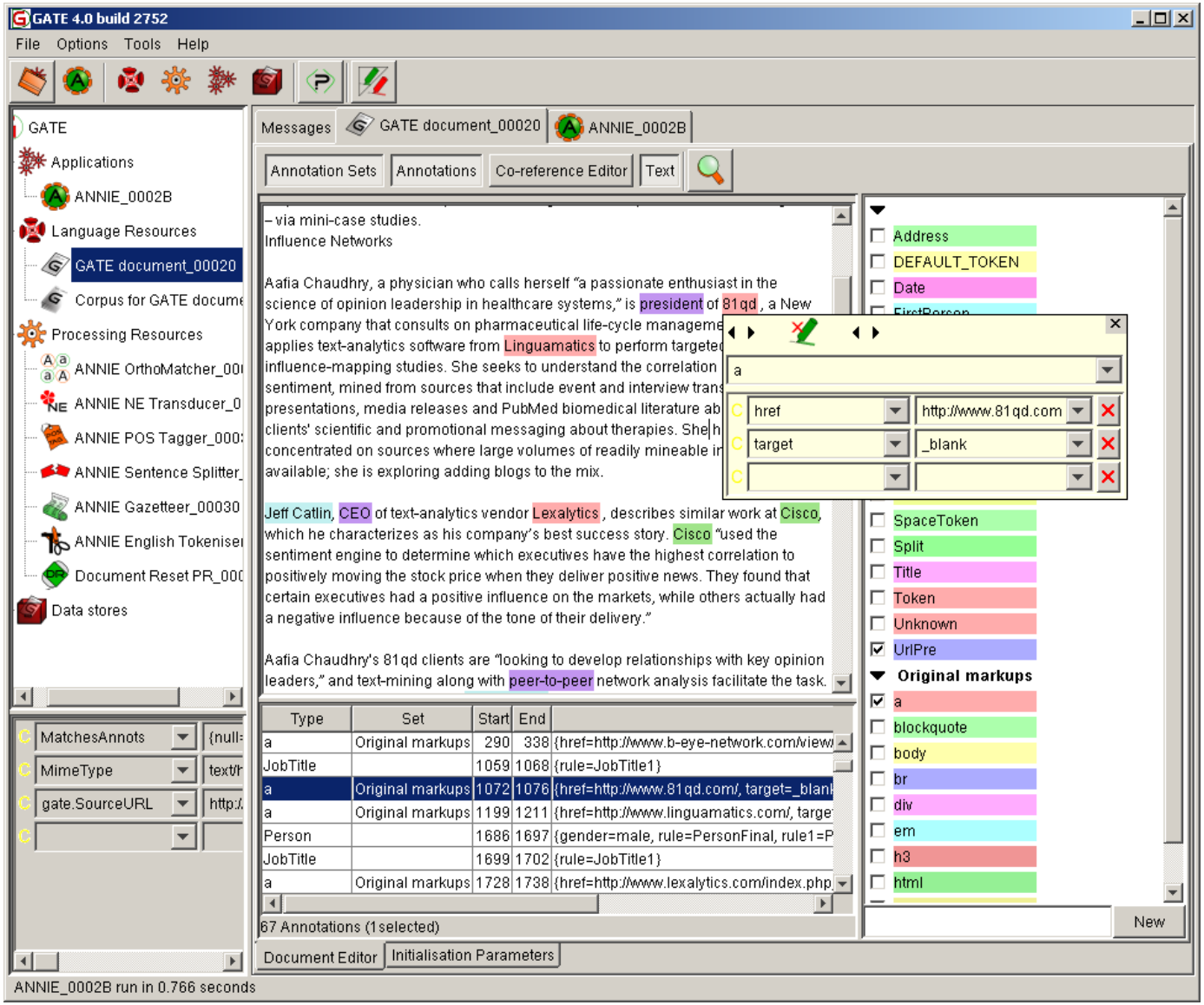

Connexor - Technology - Machinese - Demo - Machinese Syntax - demo - Mozilla Firefox
File Edit View History Bookmarks Tools Help del.icio.us
http://www.connexor.eu/technology/machinese/demo/syntax/
Google
connexor
natural knowledge
Sitemap
Home
Company
Solutions
Technology
Partners
Contact
Technology > Machinese > Demo > Machinese Syntax - demo
Machinese
Machinese Metadata
Machinese Syntax
Machinese Semantics
Machinese Phrase Tagger
Demo
Machinese Syntax
Machinese Syntax is a syntactic parser that returns base forms and compound structure, produces part-of-speech classes, inflectional tags, noun phrase markers and syntactic dependencies. Syntactic dependencies show functional relations between words and phrases in sentences.
What's the best price for new laptop that I'll use for business trips and around the office?
English text
Apply Syntax
This demo is intended for evaluation purposes only.
Done


Connexor - Technology - Machinese - Demo - Machinese Syntax - demo - Mozilla Firefox
File Edit View History Bookmarks Tools Help del.icio.us
http://www.connexor.eu/technology/machinese/demo/syntax/
connexor
natural knowledge
Sitemap
Home Company Solutions Technology Partners Contact
Technology > Machinese > Demo > Machinese Syntax - demo
Machinese
Machinese Metadata
Machinese Syntax
Machinese Semantics
Machinese Phrase Tagger
Demo
Analysis of Machinese Syntax for English:
root main: 's subj: What comp: price attr: best det: the mod: for pcomp: laptop attr: new mod: use v-ch: 'll subj: I obj: that ha: for mp: trips cc: and cc: around pcomp: office attr: business det: the
Note: The Connexor Machinese demos are intended for evaluation purposes only.
Connexor Oy, Helsinki Business and Science Park. Finland. info@connexor.com
© Connexor Oy. Powered by TomiSait
Applet Dtree started


Connexor - Technology - Machinese - Demo - Machinese Phrase Tagger - demo - Mozilla Firefox
File Edit View History Bookmarks Tools Help del.icio.us
http://www.connexor.eu/technology/machinese/demo/tagger/
connexor
natural knowledge
Sitemap
Home Company Solutions Technology Partners Contact
Technology > Machinese > Demo > Machinese Phrase Tagger - demo
Machinese
Machinese Metadata
Machinese Syntax
Machinese Semantics
Machinese Phrase Tagger
Demo
English Machinese Phrase Tagger 4.6 analysis:
Text Baseform Phrase syntax and part-of-speech
What what nominal head, pro-nominal
's be main verb, indicative present
the the premodifier, determiner
best good premodifier, superlative adjective, noun phrase begins
price price nominal head, noun, noun phrase continues
for for postmodifier, preposition, noun phrase continues
new new premodifier, adjective, noun phrase continues
laptop lap top nominal head, noun, noun phrase ends
that that nominal head, pro-nominal
I I nominal head, pro-nominal
'll will auxiliary verb, indicative present
use use main verb, infinitive
for for preposed marker, preposition
business business premodifier, noun, noun phrase begins
trips trip nominal head, plural noun, noun phrase ends
Connexor Oy, Helsinki Business and Science Park. Finland. info@connexor.com
© Connexor Oy. Powered by ToimiSait
Done


GATE 4.0 build 2752
File Options Tools Help
GATE
Applications
ANNIE_0002B
Language Resources
GATE document_00020
Corpus for GATE docume
Processing Resources
ANNIE OrthoMatcher_00
ANNIE NE Transducer_0
ANNIE POS Tagger_000
ANNIE Sentence Splitter_
ANNIE Gazetteer_00030
ANNIE English Tokeniser
Document Reset PR_000
Data stores
Messages
GATE document_00020
Annotation Sets
Annotations
Co-reference Editor
Text
Sentiment Analysis: A Focus on Applications
by Seth Grimes
Published: February 19, 2008
Text analytics can be applied to extract and analyze attitudinal information from sources as varied as articles, blog postings, e-mail, call-center notes and survey responses.
Last month, I looked at Sentiment Analysis: Opportunities and Challenges , promising a follow-on focus on applications. It's the breadth of opportunities – promising ways text analytics can be applied to extract and analyze attitudinal information from sources as varied as articles, blog postings, e-mail, call-center notes and survey responses – and the difficulty of the technical challenges that make existing and emerging applications so interesting.
We will explore three applications – influence networks, assessment of marketing response and customer experience management/enterprise feedback management – via mini-case studies.
Influence Networks
Aafia Chaudhry, a physician who calls herself "a passionate enthusiast in the science of opinion leadership in healthcare systems," is president of 81qd , a New York company that consults on pharmaceutical life-cycle management. Chaudhry
Original markups
a
blockquote
body
br
div
em
h3
html
p
span
Type Set Start End Features
a Original markups 48 59 {href=/channels/index.php?filter_channel=1394,
a Original markups 266 266 {href=http://www.clarabridge.com/, isEmptyAndSp
a Original markups 290 338 {href=http://www.b-eye-network.com/view/6744,
a Original markups 1072 1076 {href=http://www.81qd.com/, target=_blank}
a Original markups 1199 1211 {href=http://www.linguamatics.com/, target=_blan
a Original markups 1728 1738 {href=http://www.lexalytics.com/index.php, target=
a Original markups 3919 3937 {href=http://www.andersonanalytics.com/, target=
15 Annotations (1 selected)
New
MimeType
gate.SourceURL
Document Editor
Initialisation Parameters
Views built!


GATE 4.0 build 2752
File Options Tools Help
GATE
Applications
ANNIE_0002B
Language Resources
GATE document_00020
Corpus for GATE docume
Processing Resources
ANNIE OrthoMatcher_00(
ANNIE NE Transducer_0
ANNIE POS Tagger_000
ANNIE Sentence Splitter_
ANNIE Gazetteer_00030
ANNIE English Tokeniser
Document Reset PR_000
Data stores
Messages
GATE document_00020
ANNIE_0002B
Loaded Processing resources
Name Type
Selected Processing resources
! Name Type
Document Reset PR_0002C Document Reset PR
ANNIE English Tokeniser_0002D ANNIE English Tokeniser
ANNIE Gazetteer_00030 ANNIE Gazetteer
ANNIE Sentence Splitter_00031 ANNIE Sentence Splitter
ANNIE POS Tagger_00034 ANNIE POS Tagger
ANNIE NE Transducer_00035 ANNIE NE Transducer
ANNIE OrthoMatcher_00036 ANNIE OrthoMatcher
Corpus: Corpus for GATE document_00020
The corpus and document parameters are not available as they are automatically set by the controller!
No selected processing resource
Name Type Required Value
Run
Serial Application editor
Initialisation Parameters
ANNIE_0002B run in 0.766 seconds


GATE 4.0 build 2752
File Options Tools Help
GATE
Applications
ANNIE_0002B
Language Resources
GATE document_00020
Corpus for GATE docume
Processing Resources
ANNIE OrthoMatcher_00
ANNIE NE Transducer_0
ANNIE POS Tagger_000
ANNIE Sentence Splitter_
ANNIE Gazetteer_00030
ANNIE English Tokenise
Document Reset PR_000
Data stores
Messages
GATE document_00020
ANNIE_0002B
Annotation Sets
Annotations
Co-reference Editor
Text
– via mini-case studies.
Influence Networks
Aafia Chaudhry, a physician who calls herself "a passionate enthusiast in the science of opinion leadership in healthcare systems," is president of 81qd , a New York company that consults on pharmaceutical life-cycle manageme applies text-analytics software from Linguamatics to perform targete influence-mapping studies. She seeks to understand the correlation sentiment, mined from sources that include event and interview trans presentations, media releases and PubMed biomedical literature ab clients' scientific and promotional messaging about therapies. She h concentrated on sources where large volumes of readily mineable ir available; she is exploring adding blogs to the mix.
Jeff Catlin, CEO of text-analytics vendor Lexalytics , describes similar work at Cisco, which he characterizes as his company's best success story. Cisco "used the sentiment engine to determine which executives have the highest correlation to positively moving the stock price when they deliver positive news. They found that certain executives had a positive influence on the markets, while others actually had a negative influence because of the tone of their delivery."
Aafia Chaudhry's 81qd clients are "looking to develop relationships with key opinion leaders," and text-mining along with peer-to-peer network analysis facilitate the task.
Type Set Start End
a Original markups 290 338 {href=http://www.b-eye-network.com/view
JobTitle 1059 1068 {rule=JobTitle1}
a Original markups 1072 1076 {href=http://www.81qd.com/, target=_blank
a Original markups 1199 1211 {href=http://www.linguamatics.com/, targe
Person 1686 1697 {gender=male, rule=PersonFinal, rule1=P
JobTitle 1699 1702 {rule=JobTitle1}
a Original markups 1728 1738 {href=http://www.lexalytics.com/index.php
67 Annotations (1 selected)
Document Editor
Initialisation Parameters
ANNIE_0002B run in 0.766 seconds
MatchesAnnots
MimeType
gate.SourceURL
a
href
http://www.81qd.com
target
_blank
Address
DEFAULT_TOKEN
Date
SpaceToken
Split
Title
Token
Unknown
UrlPre
Original markups
a
blockquote
body
br
div
em
h3
html
New

Text Analytics

Text analytics automates what researchers, writers, scholars, and all the rest of us have been doing for years. Text analytics --

Applies linguistic and/or statistical techniques to extract concepts and patterns *that can be applied to categorize and classify documents, audio, video, images.*

Transforms "unstructured" information into data *for application of traditional analysis techniques.*

Unlocks meaning and relationships *in large volumes of information that were previously unprocessable by computer.*

*... **models text**.*

Predictive Text Analytics?

Let's consider three interpretations:

Predictive text analytics

Prediction applied to text.

Predictive analytics from text sources

Analysis of information extracted from text.

Predictive ***text analytics***

Clustering and classification of the text at document & feature levels.

Predictive Text Analytics

Predictive text analytics

Prediction applied to text.

Predictive analytics from text sources

Analysis of information extracted from text.

Predictive ***text analytics***

Clustering and classification of the text at document & feature levels.

Predictive Text

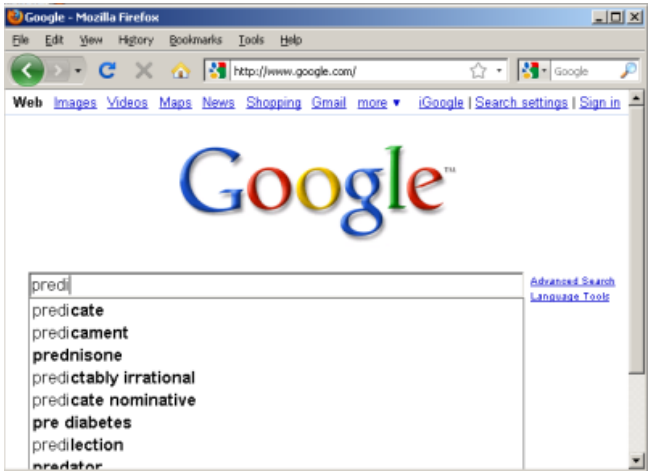
Basic modelling to facilitate functions such as:

Completion

http://en.wikipedia.org/wiki/File:ITap_on_Motorola_C350.jpg

Disambiguation

: use dictionaries, context

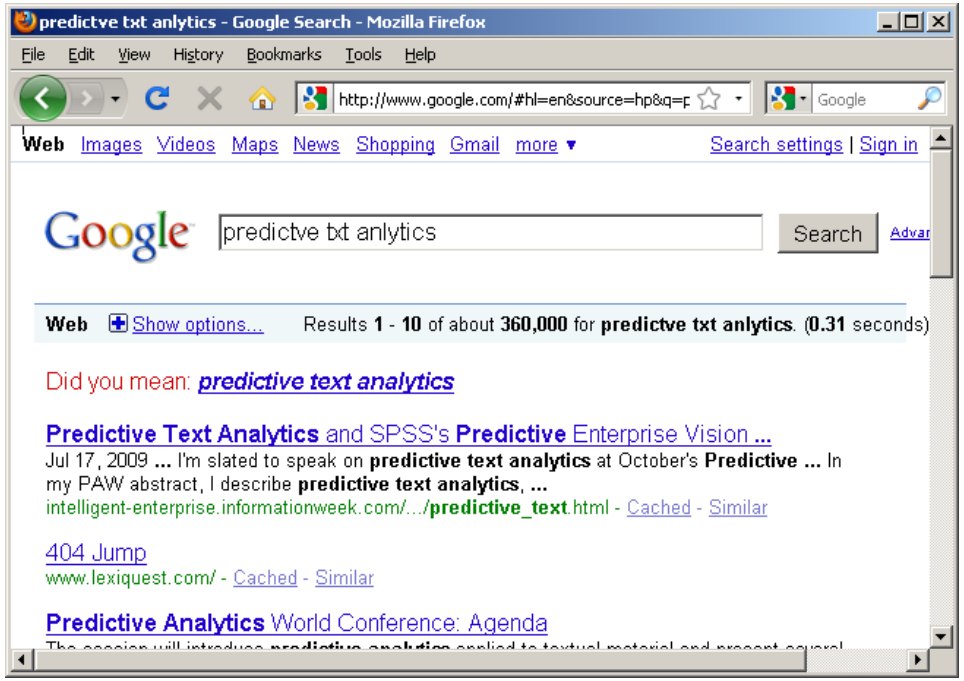


Error correction

Predictive Text

Marti Hearst in Search User Interfaces:

"Search logs suggest that from 10-15% of queries contain spelling or typographical errors. Fittingly, one important query reformulation tool is spelling suggestions or corrections."



Predictive Text Analytics

Predictive text analytics

Prediction applied to text.

Predictive analytics from text sources

Analysis of information extracted from text.

Predictive ***text analytics***

Clustering and classification of the text at document & feature levels.

Predictive Analytics, Text Sources

"The bulk of information value is perceived as coming from data in relational tables. The reason is that data that is structured is easy to mine and analyze."

– Prabhakar Raghavan, Yahoo Research, former CTO of enterprise-search vendor Verity (now part of Autonomy), 2004

Yet it's a truism that 80% of enterprise information is in "unstructured" form.

Sources

Consider:

E-mail, news & blog articles, microblogging, forum postings, and other social media.

Contact-center notes and transcripts; recorded conversations.

Surveys, feedback forms, warranty claims, case reports.

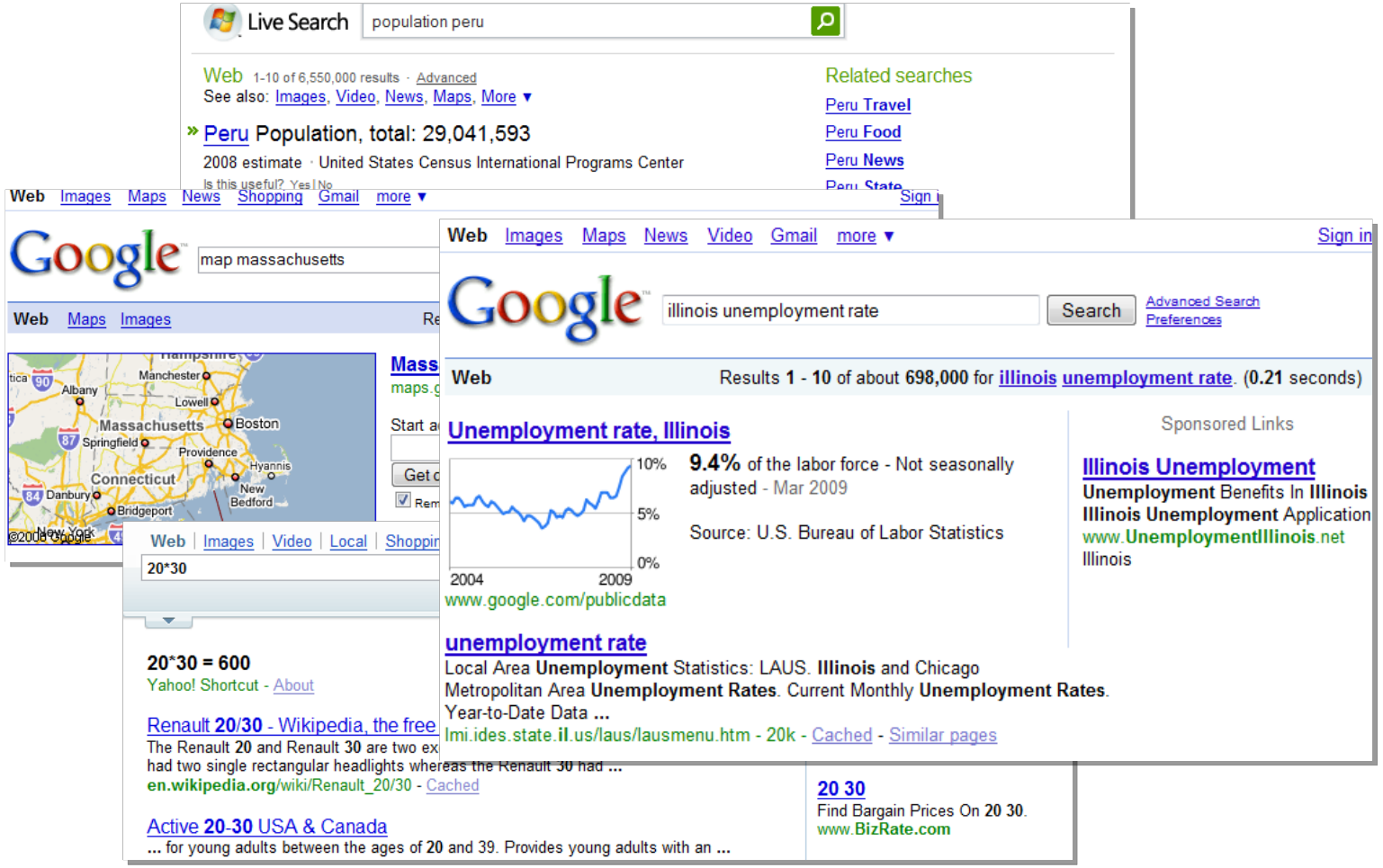
And every other sort of document imaginable.

These sources may contain "traditional" data.

The Dow fell 46.58, or 0.42 percent, to 11,002.14. The Standard & Poor's 500 index fell 1.44, or 0.11 percent, to 1,263.85, and the Nasdaq composite gained 6.84, or 0.32 percent, to 2,162.78.

Search++

Search is typically answer #1.



Beyond Search

Text analytics extracts and classifies by –

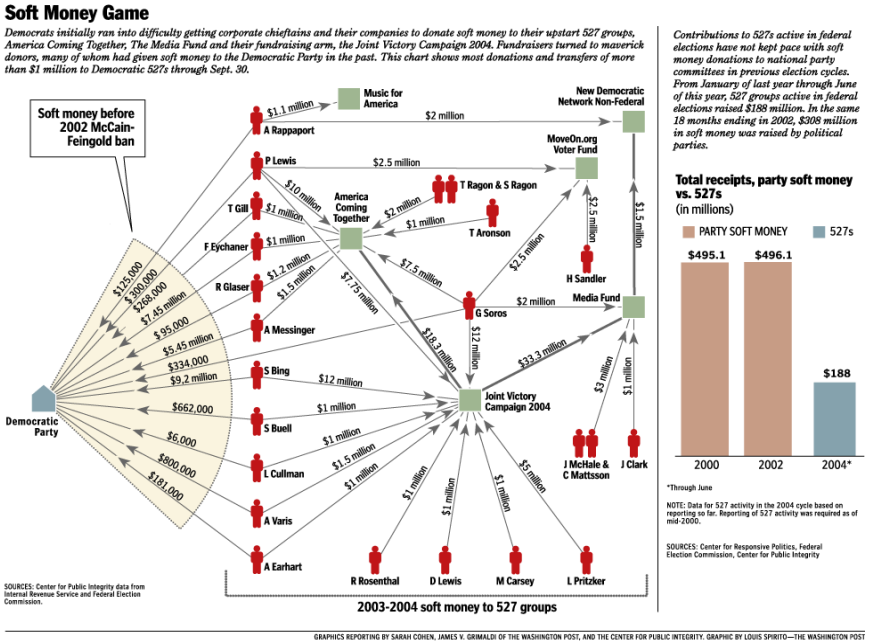
- Entities: names, e-mail addresses, phone numbers
- Concepts: abstractions of entities.
- Facts and relationships.
- Events.
- Abstract attributes, e.g., “expensive,” “comfortable”
- Opinions, sentiments: attitudinal data.

– for search indexes, knowledge bases, and databases.

Knowledge Discovery

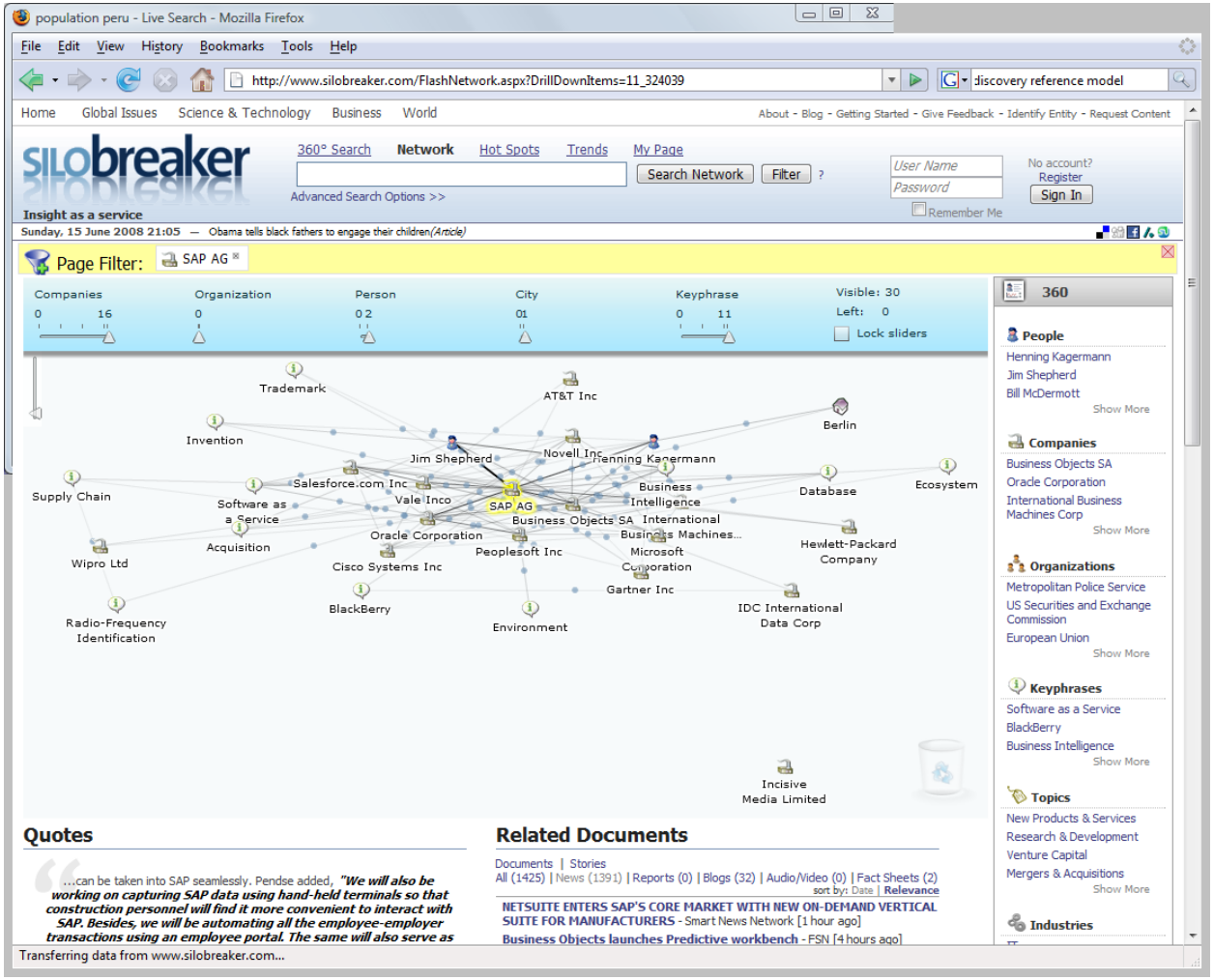
Text Mining = Data Mining of textual sources.

- Clustering and classification.
- Link Analysis.
- Prediction.
- Association rules.
- ~~Regression.~~
- ~~Forecasting~~.



Text Mining = Knowledge Discovery in Text.

Visualizing Interrelationships



Applications

Text analytics has applications in –

- Intelligence & law enforcement.
- Life sciences.
- Media & publishing including social-media analysis.
- Competitive intelligence.
- Voice of the Customer: CRM, product management & marketing.
- Legal, tax & regulatory, compliance.
- HR & recruiting.

Great *lift* potential when coupled with transactional & operational data.

Predictive Text Analytics

Predictive text analytics

Prediction applied to text.

Predictive analytics from text sources

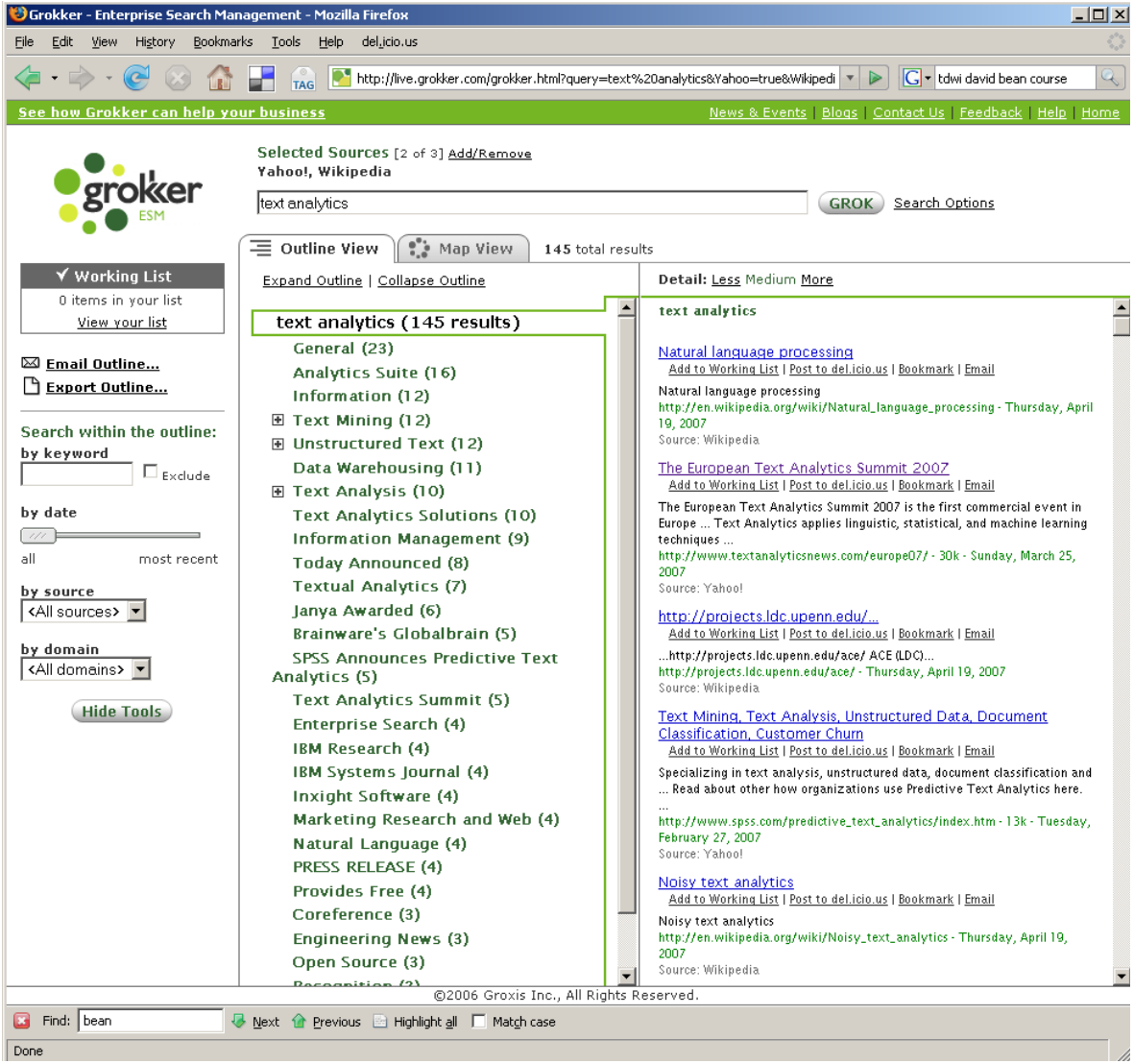
Analysis of information extracted from text.

Predictive ***text analytics***

Clustering and classification of the text at document & feature levels.

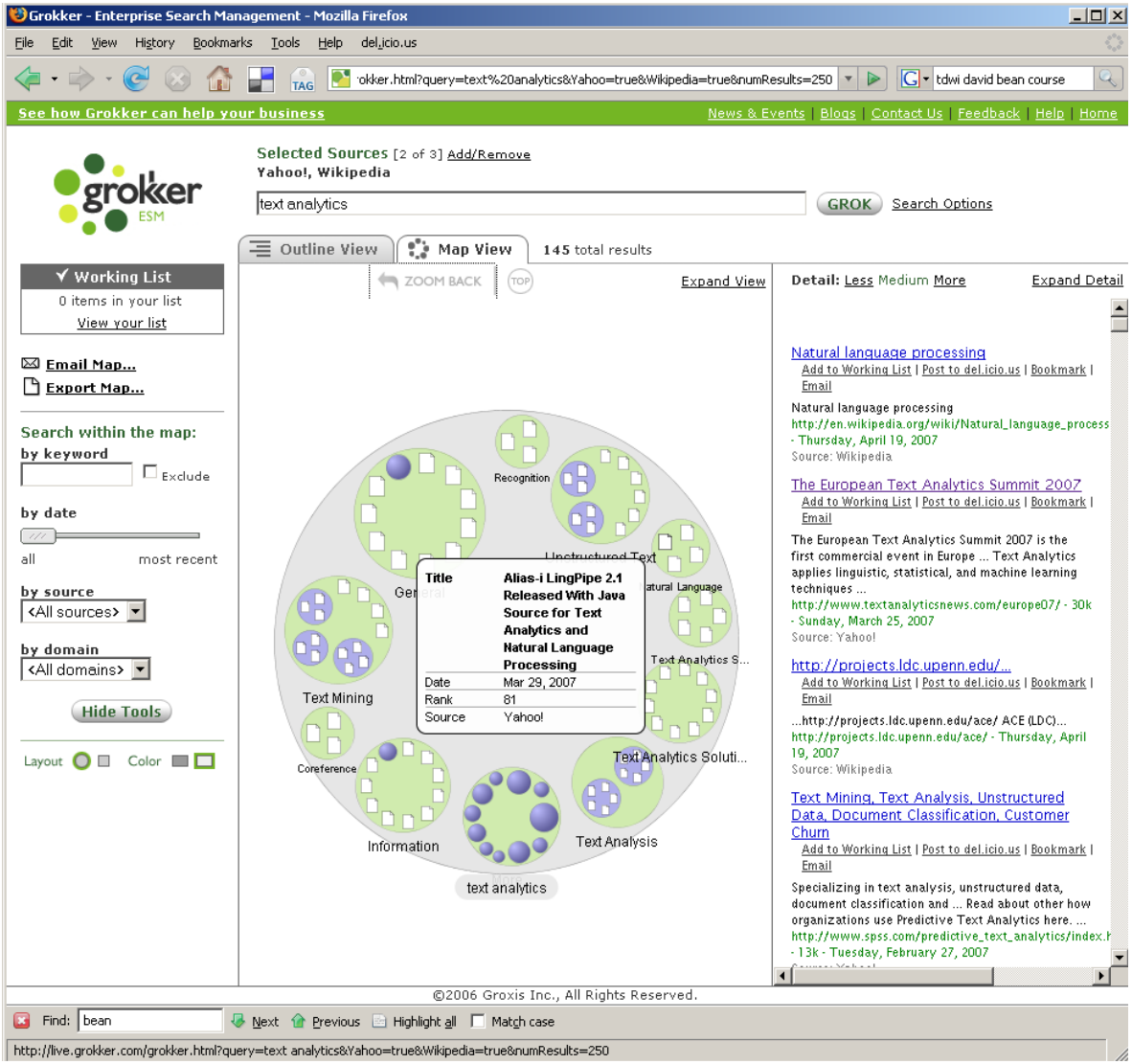
Document processing --

This slide and the next show dynamic, clustered search results from Grokker (now gone)…



…with a zoomable display.

Clustering here identifies cohesive groupings of retrieved documents.




It's Easy to Eat Right and That's a Fact! - Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://www.diet-blog.com/share/2009/04/its-e
Google
dietblog
Eat Right. Get Healthy.
Home
Share
Share » Weight Loss
It's Easy to Eat Right and That's a Fact!
By ej on Apr 15, 2009
13
0 Votes
I have read a lot of the posts and advice on here and people are saying it's all very hard to determine what to eat and what not to eat. It isn't hard - it's very easy.
If the food has been processed in any way, don't eat it. We don't need to eat a majority of the foods that we slam down our throats every day. We don't need juice, dairy, chips, sweets, ice cream etc.
People think that juice is healthy. They say there is a lot of sugar in fruit so it's bad for you. NO. Fruit has natural sugar, yes, but it also has fiber and nutrients in it that

Sentiment Analysis

"Sentiment analysis is the task of identifying positive and negative opinions, emotions, and evaluations."

-- Wilson, Wiebe & Hoffman, 2005, "Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis"

Steps include: 1) detection, 2) classification, 3) measurement:

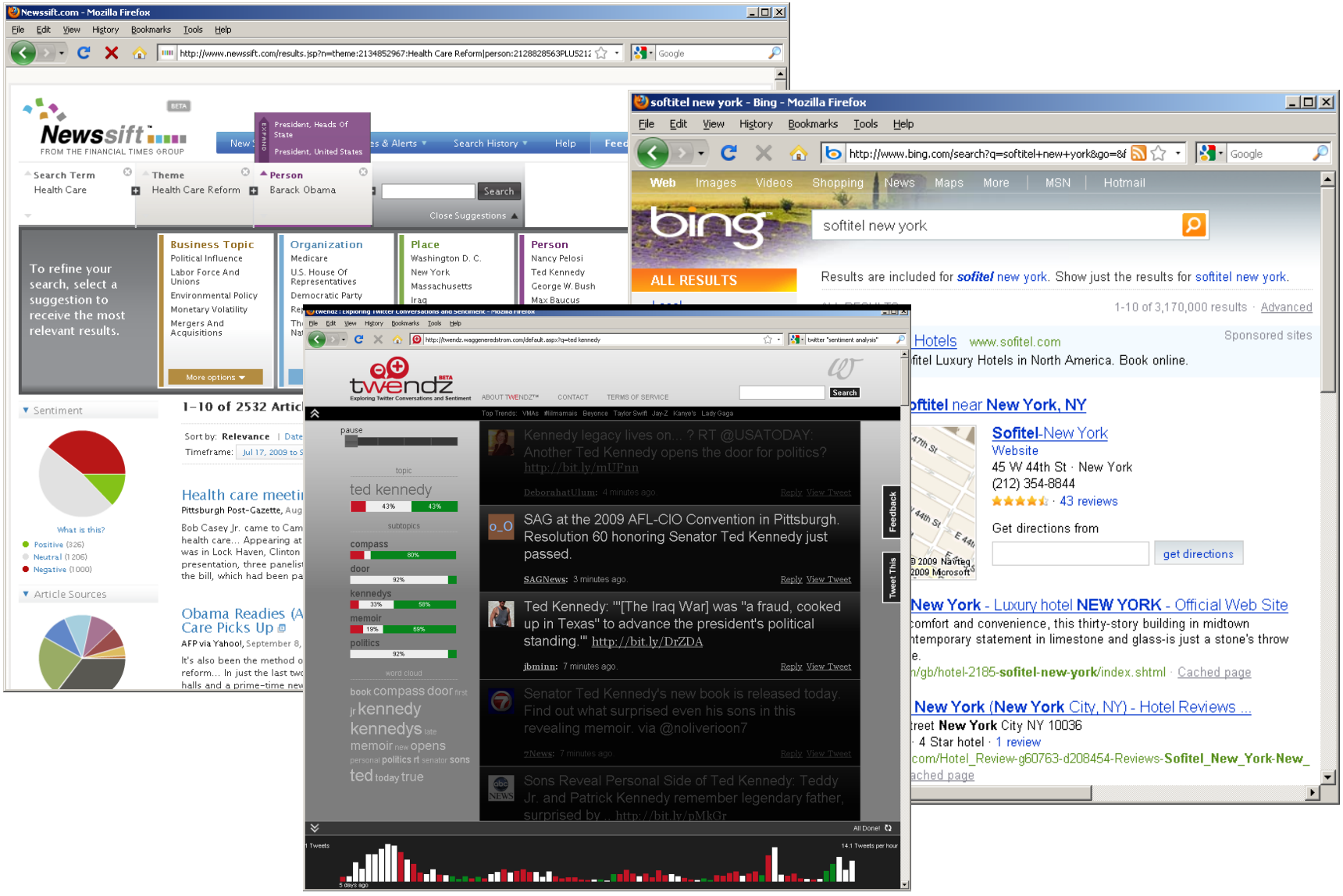
1. WW&H (for example) used over 8,000 subjectivity indicators.
2. Polarity: positive, negative, (both,) or neutral.
3. Intensity.

Sentiment Analysis

Complications.

- Levels:
  - Corpus / data space, i.e., across multiple sources.
  - Document.
  - Statement / sentence.
  - Entity / topic / concept.
- Language characteristics:
  - Jargon, slang, irony, ambiguity, anaphora, polysemy, synonymy, etc.
  - Context is key. Discourse analysis comes into play.
- Sentiment holder ≠ object:
  - *Geithner said unemployment will worsen…*

Steps in the Right Direction



Jodange | Top of Mind Service

Welcome, demo | Log out

Keyword: alternative energy (271 Total Opinions)

alternative energy Search

Home

Search results can be viewed in List, Chart or Heatmap views. Search Result Filters, available on the left, provide you with live filters to show you only the results you want to see. (hide)

List View | Opinion Index | Doppler View

Showing only opinions matching:
(remove all)

Topic: Alternative Energies, Urban Planning/Development

Opinion Holder: International Energy Agency, Boone Pickens, Energy Information Administration

Create an alert

Look Back: 1d 1w 1m 3m 6m

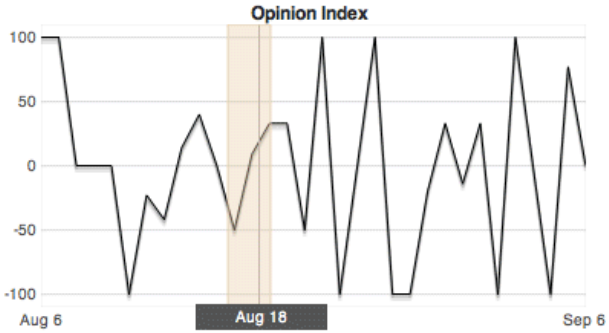
Opinion Holder (4440)
clear selected | show all

- [ ] Barack Obama (331)
- [ ] John McCain (305)
- [x] International Energy Agency (132)
- [ ] Nancy Pelosi (78)
- [ ] Sarah Palin (76)
- [x] Boone Pickens (65)
- [ ] Congress (62)
- [x] Energy Information Administration (53)
- [ ] Energy Department (47)
- [ ] George W Bush (44)
- [ ] BRITISH ENERGY GROUP PLC (33)
- [ ] International Atomic Energy Agency (33)
- [ ] Gordon Brown (31)
- [ ] Government (31)
- [ ] Reuters (31)
- [ ] Angelo Reyes (29)
- [ ] Department of Energy (29)
- [ ] Bloomberg (27)
- [ ] Vladimir Putin (27)
- [ ] U.S. Department of Energy (26)

Topic (3357)
clear selected | show all

- [x] Alternative Energies (8682)

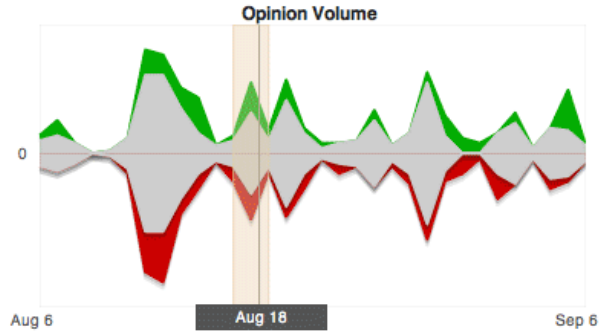
August 06, 2008 - September 06, 2008



Show Aggregate JOI

	JOI	9%		

Add Ticker Symbol Add Reset



Total	Pos	Neu	Neg
27	6	16	5

Opinion Summary
Aug 17 - Aug 19

By **Boone Pickens** in Herald-Sun on 2008-08-19

Since he formed Pickens Fuel Corp in 1997, Pickens has been arguing that gas is the best alternative fuel for cars.

Topics: Alternative Energies, Analyst Comment/Recommendation, National/Presidential Elections, Urban Planning/Development

By **International Energy Agency** in Energy Bulletin on 2008-08-19

After assessing all aspects of the situation, the International Energy Agency warned last week that while oil consumption in the US is expected to fall by 3.1 percent this year and 2 percent next year, Chinese oil consumption, while only a third that of the US, is expected to grow by about 5.6-5.7 percent in the next two years, thereby offsetting much of the drop in US consumption.

Topics: Alternative Energies, Climate Change

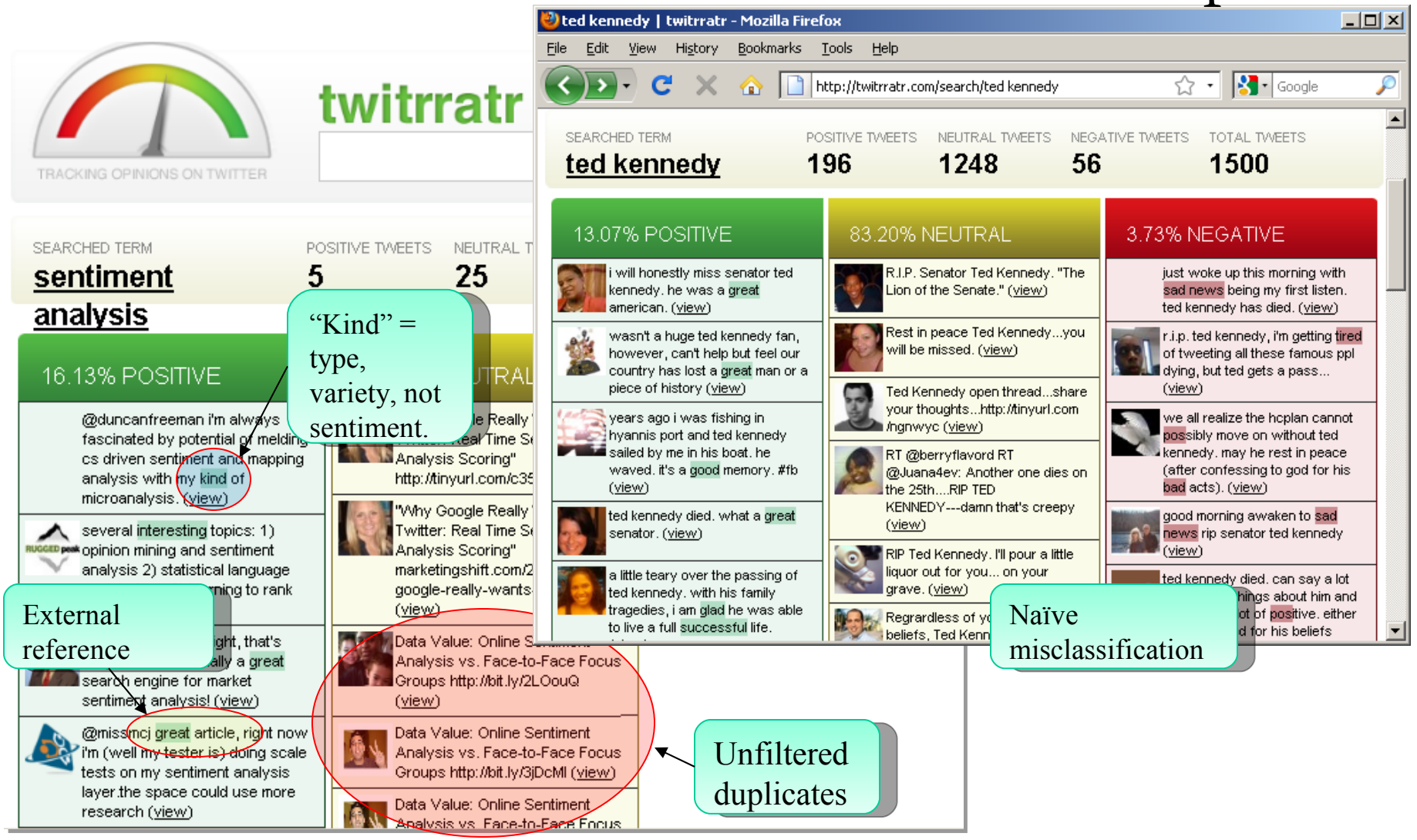
By **Energy Information Administration** in Pollution Online on 2008-08-19

In May 2008 the Energy Information Administration, which provides official energy statistics from the U S Government, stated that in 2004 the United States produced about 22 percent of global carbon dioxideemissions, primarily because the U Secnomy is the

Correlate the Opinion Index and Volumes to stock price changes

CLOSE

... and Missteps

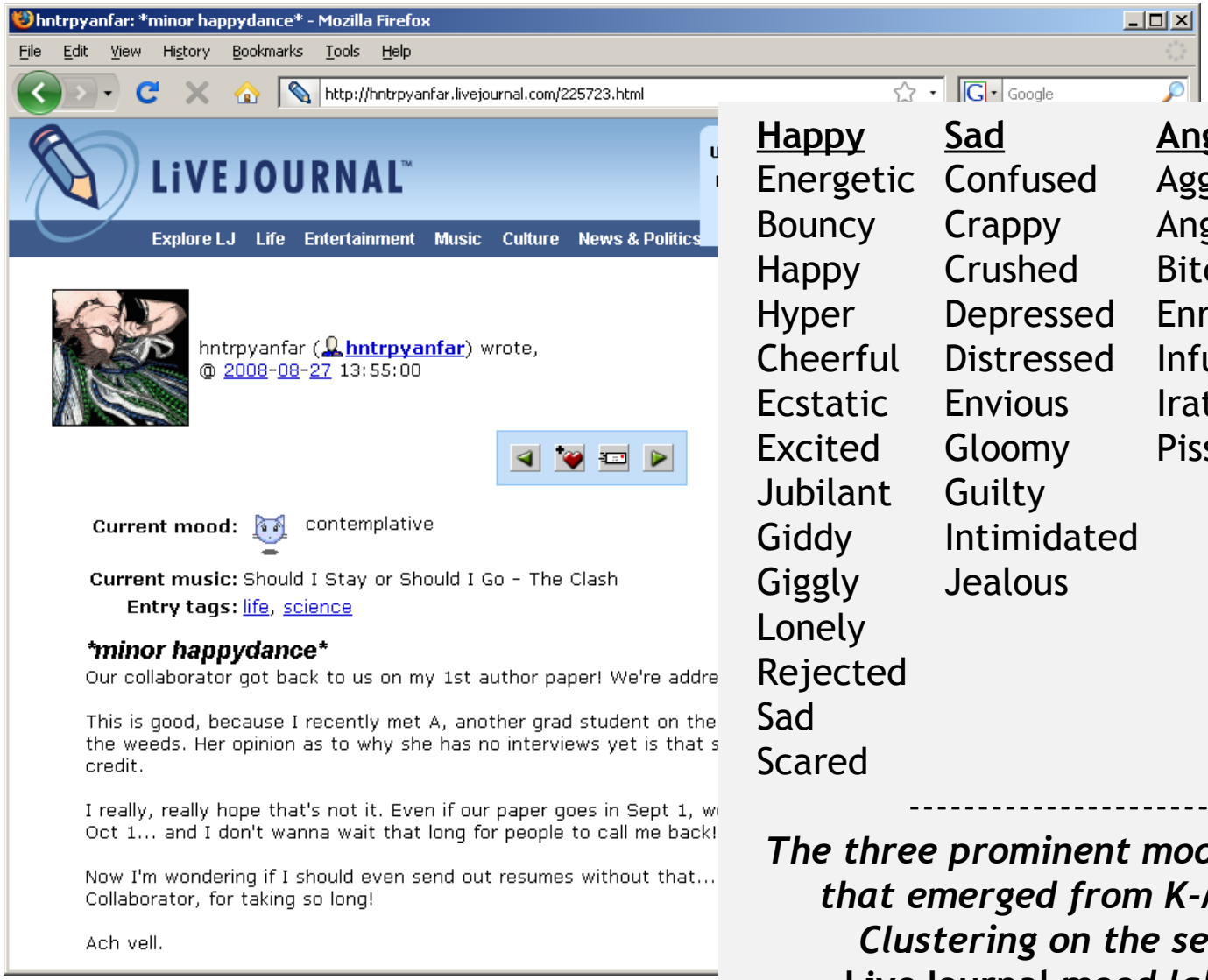


Alta Plana

Beyond Polarity

"We present a system that adds an emotional dimension to an activity that Internet users engage in frequently, search.."

-- Sood & Vasserman & Hoffman, 2009, "*ESSE*: Exploring Mood on the Web"



Happy	Sad	Angry
Energetic	Confused	Aggravated
Bouncy	Crappy	Angry
Happy	Crushed	Bitchy
Hyper	Depressed	Enraged
Cheerful	Distressed	Infuriated
Ecstatic	Envious	Irate
Excited	Gloomy	Pissed off
Jubilant	Guilty	
Giddy	Intimidated	
Giggly	Jealous	
Lonely		
Rejected		
Sad		
Scared		

The three prominent mood groups that emerged from K-Means Clustering on the set of* LiveJournal *mood labels.

Predictive Text Analytics

Predictive text analytics

Prediction applied to text.

Predictive analytics from text sources

Analysis of information extracted from text.

Predictive ***text analytics***

Clustering and classification of the text at document & feature levels.

Seth Grimes

Alta Plana Corporation

301-270-0795 – *http://altaplana.com*

@SethGrimes