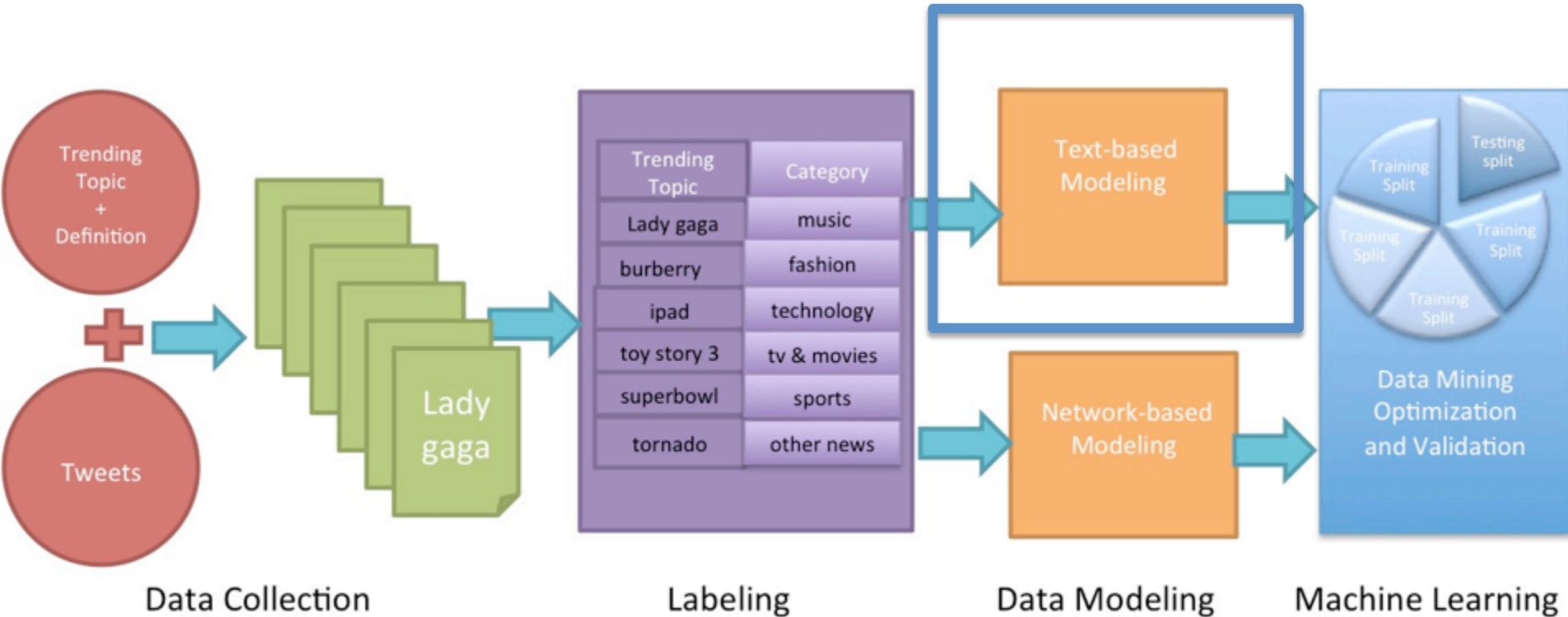


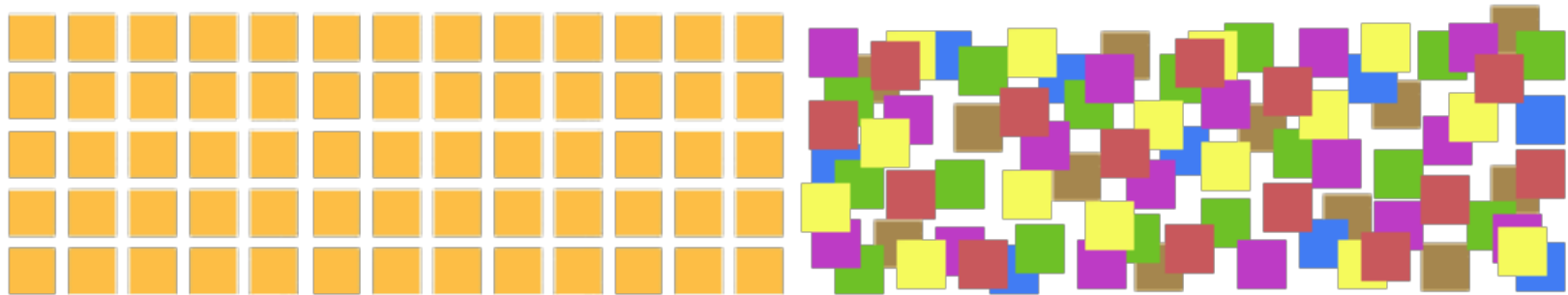
Extracting Social Network data for BI: workflow



Two types of data: network-based and **text-based**

The power of unstructured data

“80% of business-relevant information originates in unstructured form, primarily text.”

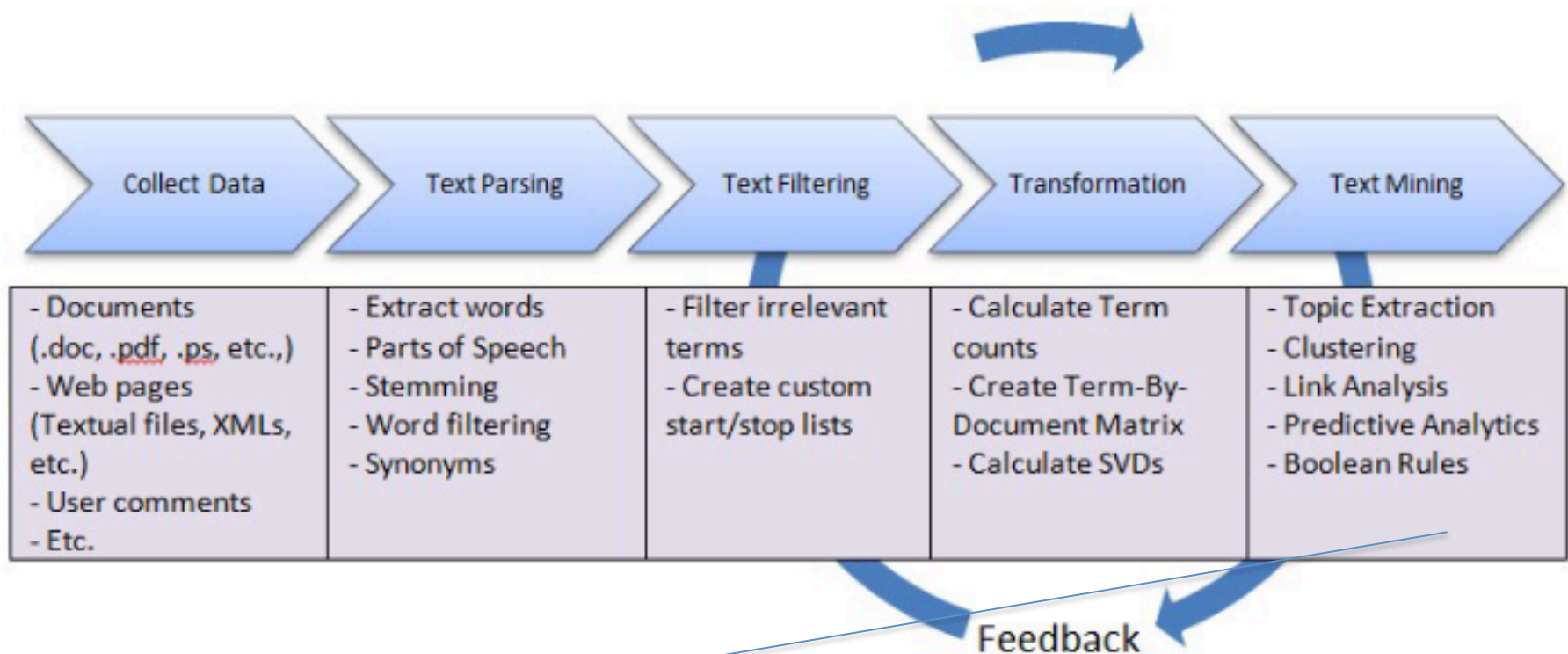


Structured Data vs. **Unstructured Data**

Text Mining

- Text mining is an emergent technology attempting to extract useful information (and knowledge) from unstructured data
- Text mining is an extension of data mining to textual data
- Social networks contain a lot of information in textual form, such as posts, links, blogs, news articles, emails..

Text processing workflow (more in detail)



TEXT MINING
APPLICATIONS

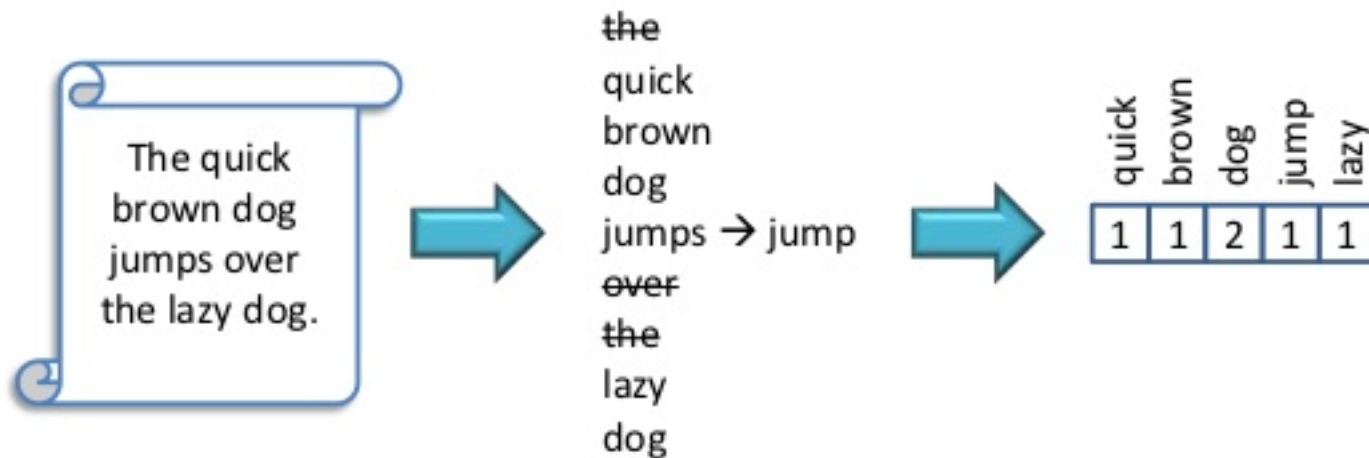
Text Mining: popular applications

- Named Entities “who/what/where?”
- Themes/Topics “what’s the buzz?”
- Categories “what’s it about?”
- Intentions “what will they do?”
- Sentiment “how do they feel?”

How do we represent texts? Bag of words

Bag of words

- Tokenize
- Remove stop words
- Lemmatize
- Compute weights



Seems simple.. however..

Of all the sensory impressions proceeding to the brain, the visual experiences are the dominant ones. Our perception of the world around us is based essentially on the messages that reach our eyes. For a long time, the retinal image was considered as a simple picture. However, a movie camera shows that an image is not a simple picture. It is a discovery that the eye is not a simple camera. We now know that the eye is a complex organ. The perception of the world is more complex than we thought. Following the work of Hubel and Wiesel, we now know that the message about the image falling on the retina undergoes a complex analysis in a system of nerve cells. The information is stored in columns. In this system each cell has its specific function and is responsible for a specific detail in the pattern of the retinal image.



**sensory, brain,
visual, perception,
retinal, cerebral cortex,
eye, cell, optical
nerve, image
Hubel, Wiesel**

China is forecasting a trade surplus of \$90bn (£51bn) to \$100bn this year, a threefold increase on 2004's \$32bn. The Commerce Ministry said the surplus would be created by a predicted 30% increase in exports to \$750bn, compared with \$575bn in 2004. Imports are expected to rise to \$660bn. The increase in exports will annoy the US. China's government has deliberately kept the yuan undervalued against the dollar. It has agreed to increase exports to the US. The yuan is undervalued against the dollar. The government also needs to increase exports to the US. The demand for the yuan is high. The country. China's government has permitted it to trade within a narrow range but the US wants the yuan to be allowed to rise freely. However, Beijing has made it clear that it will take its time and tread carefully before allowing the yuan to rise further in value.



**China, trade,
surplus, commerce,
exports, imports, US,
yuan, bank, domestic,
foreign, increase,
trade, value**

Why bag of words?

– Basic idea:

- Keywords are extracted from texts.
- These keywords describe the (usually) topical content of Web pages and other text contributions.
- Each unique word in a corpus of documents (web pages, social messages..) = one attribute
- Each document is a record with non-zero weight for each word in that document, zero weight for other words

➔ Words become “attributes”, whose values can be binary (the word *is* or **is not** in a text), or real numbers (e.g. the relative frequency of a word in the text)

Example

the dog is on the table

0	0	1	1	0	1	1	1
are	cat	dog	is	now	on	table	the


The example considers a vocabulary of 8 words – in the reality the vocabulary has millions of words – document records have millions of attributes


Which words should we care about? A complex problem

- E.g.: Companies assume that people refer to them by name
- Big mistake!!
- There are multiple dimensions for reference: hashtags, names of people (e.g., managers), products, and for each: abbreviations, initials, nicknames
- Additional problems: ambiguity, synonyms..


Example: search “Watson” on Twitter Search

The image shows a screenshot of a Twitter search for the name "Watson". At the top, two profile cards are displayed side-by-side. The left card is for Emma Watson, featuring a circular profile picture and a "Segui" button. Her bio identifies her as an actor and UN Women Global Goodwill Ambassador. The right card is for Deshaun Watson, a professional football player, also with a "Segui" button. His bio includes his team information and contact details. Below these profiles, a tweet from CommentWise (@oncommentwise) is visible, discussing a video about fictional detectives. At the bottom right, a tweet from Digital Marketing (@DollyRayDigital) is shown, featuring a video thumbnail for an IBM Watson Education video titled "Personalizing the teaching ...". The video thumbnail shows a play button icon over a background of colorful icons representing various digital and educational concepts. Below the video, there are icons for replies, retweets (91), and likes (4).

Emma Watson 
@EmmaWatson
Actor & @UN_Women Global Goodwill Ambassador. Facebook: EmmaWatson Instagram: EmmaWatson Goodreads: OurSharedShelf

Deshaun Watson 
@deshaunwatson
God 1st! •815™ •GodSpeed •Memo™ •six. #NEGU For Football Inquiries Contact: @DavidMulugheta For Marketing & Business Inquiries...

CommentWise @oncommentwise · 4 min
Evidence, my dear **Watson!**: Kajal Iyer As a country we have grown up on fictional detectives who solve cases by... divr.it/PvJqQD

Digital Marketing  @DollyRayDigital · 2 h
#IBM #Watson Education Personalizing the teaching & learning experience youtu.be/ZvGhbJ8V8eA #Cognitive #ArtificialIntelligence #AI #IoT

IBM Watson Education Personalizing the teaching ...
youtube.com

91 4

So.. rather than “bag of words”, bag of concepts

Annotation Sets Annotations Co-reference Editor Text

Background to and purpose of visit

The enquiry had come via **Jean Coldham** of The British **Midlands** and originally from the UKTI in the British Consulate in **Chicago**. UKTI had been alerted of the possible expansion project via **Sterigenics HQ** in **Chicago**, who had supplied **Ron Peacock**'s contact details. **Ron** had previously helped emda with a number of research projects, agreeing to answer questions and give background information etc.. This was the first indication, however, of a possible expansion project for the Company in **Somercotes**, part of the EM SFI grant assisted area. Local contact would be followed by a meeting with the Group FCO in the **US**, involving UKTI staff and **Jean**.

NB **Ron Peacock** originally set the Company up in **1992**, is nominally based in **European HQ** in **Belgium**, although he spends the bulk of his time in **UK**. He is now responsible for special projects and capital investments.

Company Background

The Company was set up in **1992** as a spin off of **Griffith Laboratories** (next door) as **Griffith Microscience**. In **1998** it was floated on **Nasdaq** and eventually acquired by **IBA S & I**. IBA then decided that the Company was 'non-core', and sold it in turn to **Sterigenics**, a **US** owned sterilisation company. Relatively recently the **Sterigenics Group** was acquired by the **PPM Capital**, the venture capital group.

World HQ is in **Chicago**, **European HQ** in **Leuven, Belgium** and the Group has recently built new facilities in **Germany** and **China**.

Type	Set	Start	End	Features
Location		1684	1695	{locType=[null], matches=[2982, 2955, 2970], rule1=LocationPost, rule2=LocFinal}
Location		1700	1707	{locType=[null], matches=[2956, 2971], rule1=InLoc1, rule2=LocFinal}
Location		1752	1754	{locType=[null], matches=[2977, 2979, 2957, 2974], rule1=InLoc1, rule2=LocFinal}
Organization		1836	1861	{matches=[2958, 2980, 2975, 3097, 3099], orgType=[null], rule1=OrgXEnding, rule2=OrgFi}
Date		1876	1880	{kind=date, matches=[2959, 2954], rule1=YearContext1, rule2=DateOnlyFinal}
Organization		1898	1919	{orgType=[null], rule1=OrgXBase, rule2=OrgFinal}

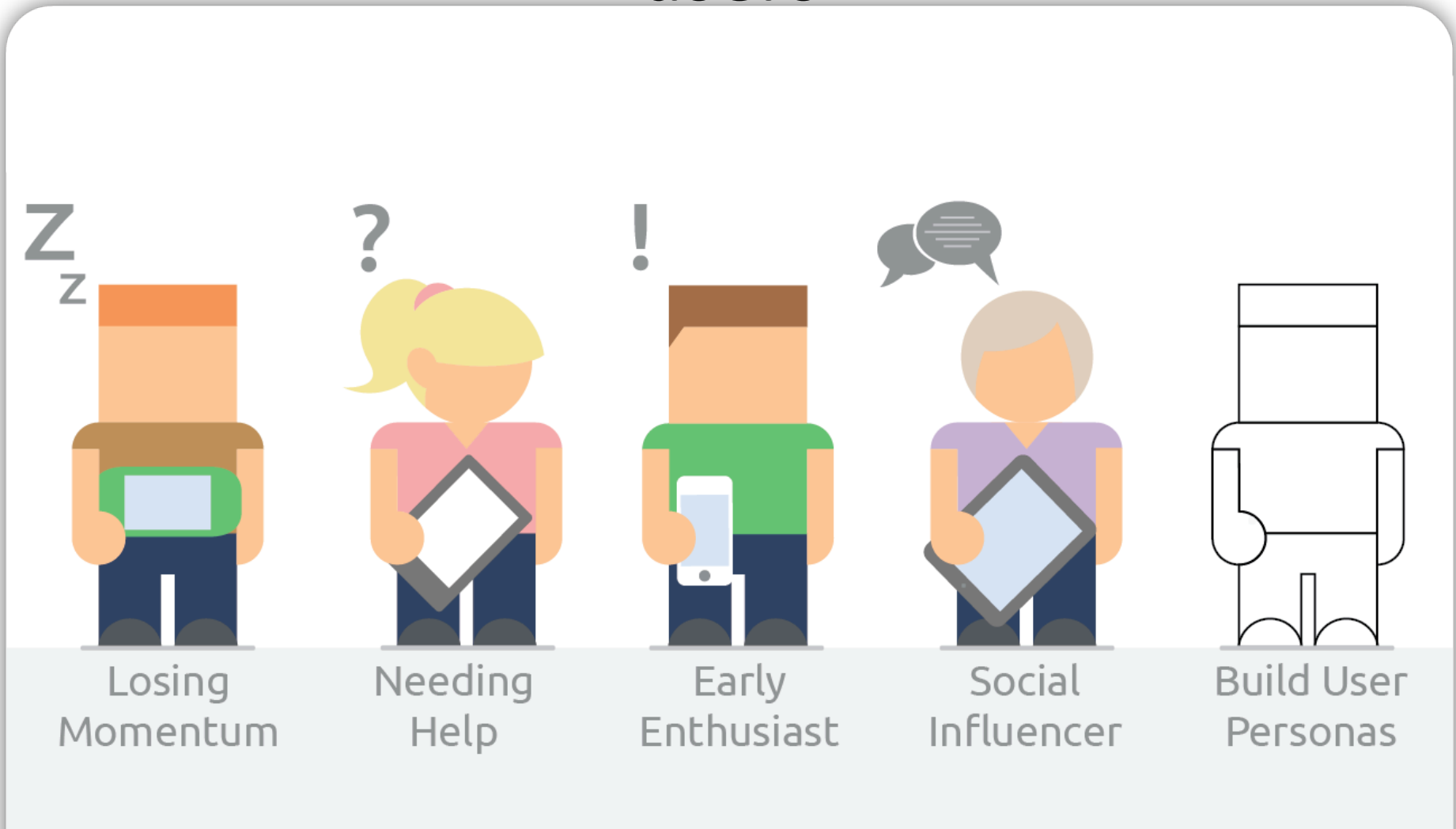
86 Annotations (1selected)
New

- Address
- DEFAULT_TOKEN
- Date
- FirstPerson
- JobTitle
- Location
- Lookup
- Money
- Organization
- Percent
- Person
- Sentence
- SpaceToken
- Split
- Temp
- TempLocation
- Title
- Token
- Unknown

Text Mining tasks

- After pre-processing, text is analyzed to extract useful information, much in the same way as for structured data mining
- Types of processing:
 - Classification: assign a message (or a piece of text) a category
 - Topic extraction: identify (summarize) what the text is about
 - Clustering: grouping similar messages/blogs/posts
 - Sentiment analysis: determine the polarity/sentiment/opinion expressed in a text

Classification example: mobile apps users



In classification, categories are assigned. Based on messages, classify user in one of these cats

Another example: Happiness

The screenshot shows a Mozilla browser window titled "happy birthday jenny!!! - Mozilla". The address bar displays "http://community.livejournal.com/birthdaybuckle/". The page content includes a message from "phineasjones" dated "Fri, Nov. 19th, 2004, 03:36 pm" with the mood tag "smurfy!". The message text is: "whreeeeee!!! it's your birthday!!! ::dances:: this makes three whole birthdays of yours that i've known you for. can you believe it? wacky. i truly hope this one is the very best of all of them. and that this year brings you still more happiness and more love and more fun - all the good things. you deserve them all, dear friend. i love you! ♥". A small profile picture of a man with glasses is visible. The "Current Mood" is "happy". A "Link" and "Leave a comment" option are present. On the left, a "Update Journal - Mozilla" window is open, showing a list of mood tags: envious, exanimate, excited, exhausted, flirty, frustrated, full, geeky, giddy, giggly, gloomy, good, grateful, groggy, grumpy, guilty, happy, high, hopeful, horny. The "Options" section includes "Security:", "Auto-Format HTML:", "Current Location:", "Music:", and "Mood:".

Since messages are tagged with a mood by the message author, a machine learning classifier can learn to classify new untagged messages

Message Clustering: example

related themes

over the last 50 minutes

#wikipedia	wolfram alpha
wolframalpha.com/	computational
fructose in onion	computational
with your input	knowledge
google killer	engine
wolfram	via @time
wolfram alpha	been alive
@time	bit.ly/sa6be
days old	got the number
meaning of life	wrong answer
i have tried	queries failed
on the day	eg
was born	i was born
engine	search engine
#fail	input

tweets by theme

“queries failed”

: not finding #WolframAlpha all I had hoped, data-driven queries failed every time ("ave # job applicants interviewed to fill position") *peterjwolfgang*

So far #WolframAlpha has failed for the two queries that I have tried that weren't a play test #fail *marcad* [show 1 similar tweet »](#)

(drilldown ↑)

“google killer”

#WolframAlpha is going to be huge, but it's not a google killer, and most people will never use it *briggs1*

@charlesf11 Yeah, it's bugging me that people are calling it a Google killer. They're not the same animal. #wolframalpha #wikipedia #google

mattbramanti

#WolframAlpha is no Google killer, but it's the smartest calculator ever.

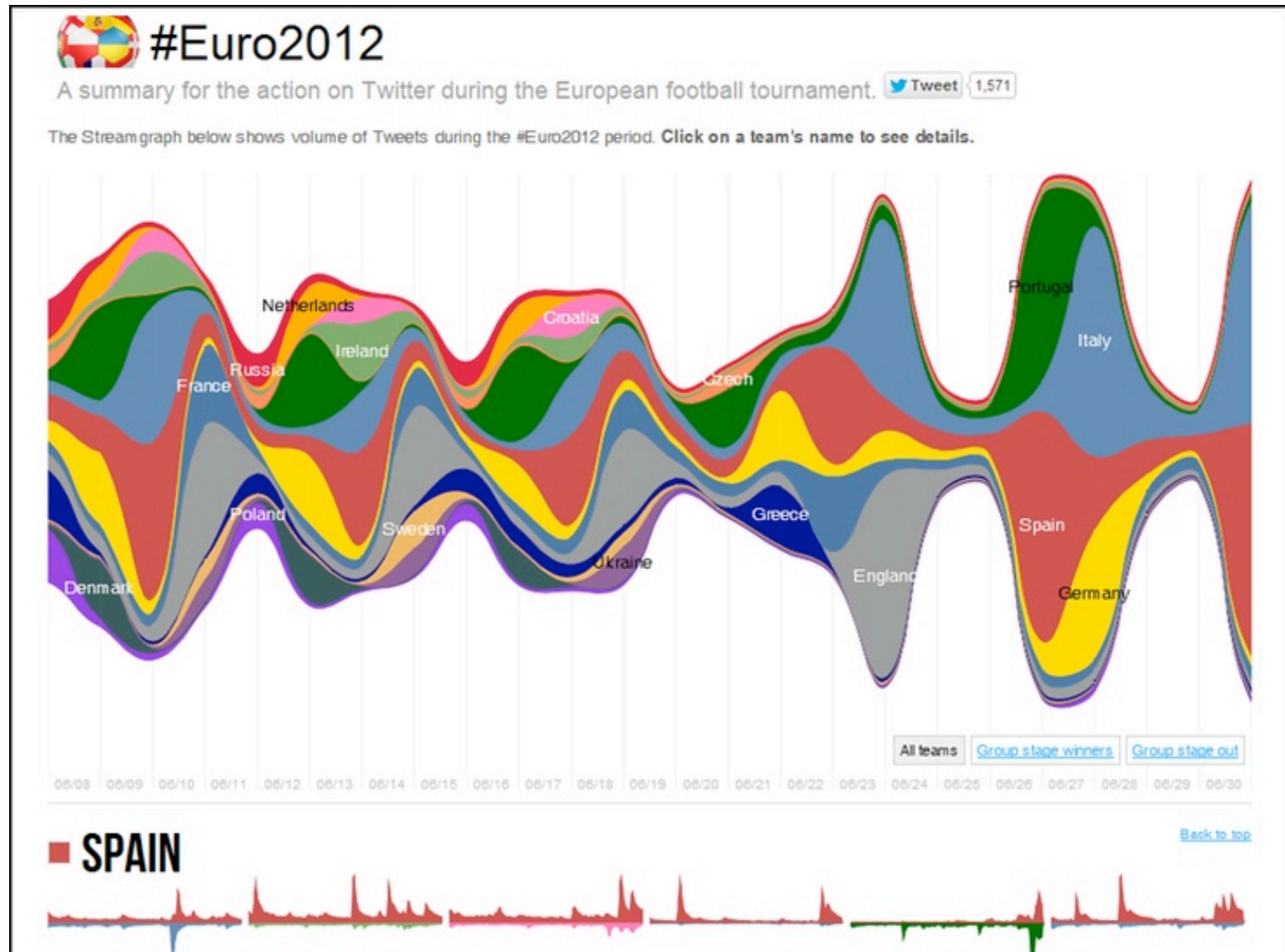
PrestonStahley

Remember: in clustering we have no known categories. Messages are clustered based on similarity of attributes (bag of words)

Topic extraction

- In topic extraction, the purpose is to detect group of words that are representative of a “trending” discussion topic
- Based on the analysis of terms that co-occur in many messages
- Example: *Watson, cognitive, IBM, AI*
- Example 2: *#WomenBoycottTwitter, women, sexual, solidarity*

Tracing topics (stream graphs)



Tracing Topics (TamTamy-Reply)



Home



Social Network



Social Buzz

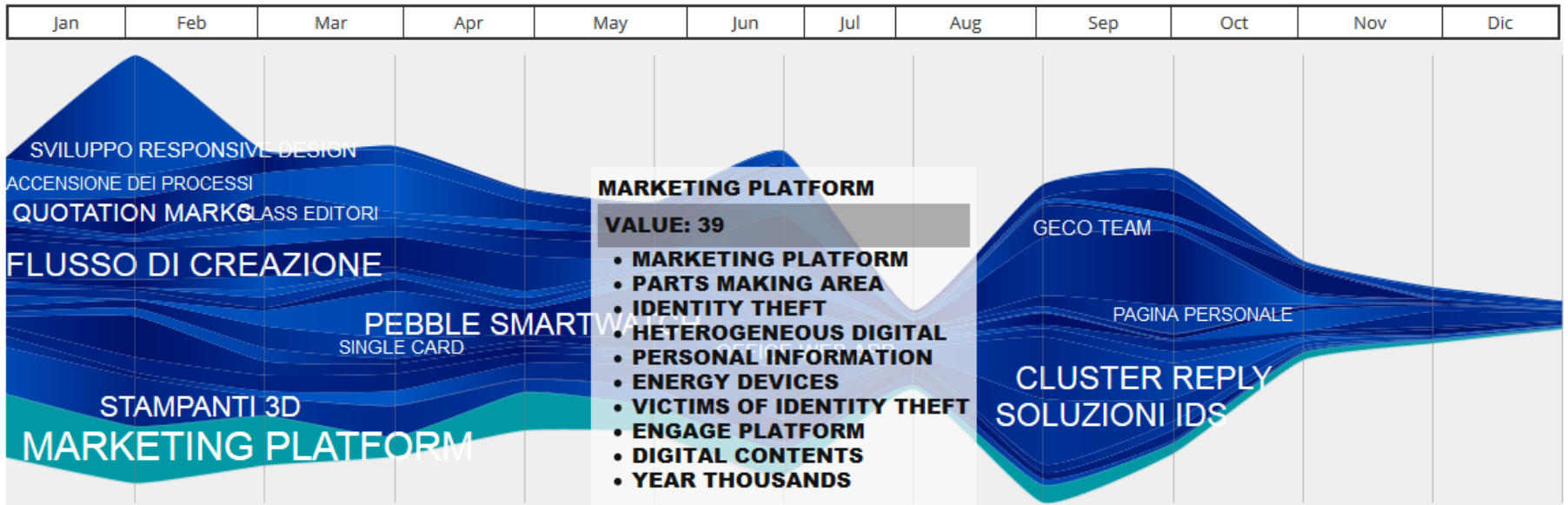


Statistics



People

⚙ Personalize Streamgraph Parameters



Opinion mining



Introduction – facts and opinions

- Two main types of textual information on the Web.
 - Facts and Opinions
- Current search engines search for facts (assume they are true)
 - Facts can be expressed with topic keywords.
- Search engines do not search for opinions
 - Opinions are hard to express with a few keywords
 - What do people think of Motorola Cell phones?
 - Current search ranking strategy is not appropriate for opinion retrieval/search.

Opinions are user-generated content

- **Word-of-mouth on the Web**
 - One can express personal experiences and opinions on almost anything, at review sites, forums, discussion groups, blogs ... (called the user generated content.)
 - They contain valuable information
 - Web/global scale!!
 - **Our interest:** to mine opinions expressed in the user-generated content
 - A very challenging problem.
 - Practically very useful.

Applications

- **Businesses and organizations:** product and service benchmarking. Market intelligence.
 - Business spends a huge amount of money to find consumer sentiments and opinions.
 - Consultants, surveys and focused groups, etc
- **Individuals:** interested in other' s opinions when
 - Purchasing a product or using a service,
 - Finding opinions on political topics,
- **Ads placements:** Placing ads in the user-generated content
 - Place an ad when one praises a product.
 - Place an ad from a competitor if one criticizes a product.
- **Opinion retrieval/search:** providing general search for opinions
 - Predicting behaviours and trends in finance, medicine, politics

Blah blah blah blah blah blah
blah. **Talking about cell
phones. I bought the new
[redacted] today. It's so
beautiful!** blah blah blah blah
blah blah blah blah blah
blah blah blah blah I **also like
[redacted]. However I cannot
recommend the [redacted]
due to its weak battery. It's
a shame.** blah blah blah
blah



Talking about cell phones. I
bought the new [redacted]
today. It's so **beautiful!**

behavior



I also like [redacted] However I **cannot
recommend** the [redacted] due
to its **weak battery.** It's a shame.

brand



feature/attribute



Impact

- 81% of Internet users have done online research on a product 20% do so on a typical day
- Among readers of online reviews between **73% and 87%** report that reviews had a significant influence on their purchase
- Consumers report being willing to pay from **20% to 99%** more for a 5-star-rated item than a 4-star-rated item (the variance stems from what type of item or service is considered);
- 32% have provided a rating on a product, service, or person via an online ratings system, and 30% have posted an online comment or review regarding a product or service.

A formalization of the opinion mining task

- Basic components of an **opinion**:
 - **Opinion holder**: The person or organization that holds a specific opinion on a particular object.
 - **Object**: on which an opinion is expressed (it can be described by features, e.g. for an hotel room: dimension, clean, silent, cost,..)
 - **Opinion**: a view, attitude, or appraisal on an object (or object feature) from an opinion holder.



Opinion mining “grain”

- At the **document (or review) level**:
 - Task: sentiment classification of reviews
 - Classes: positive, negative, and neutral
 - Assumption: each document (or review) focuses **on a single object** (not true in many discussion posts) and contains opinion from a single opinion holder.
 - Example: Movie reviews
- At the **sentence level**:
 - Task 1: identifying subjective/opinionated sentences
 - Classes: objective and subjective (opinionated)
 - Task 2: sentiment classification of sentences
 - Classes: positive, negative and neutral.
 - Assumption: **a sentence contains only one opinion**; not true in many cases.
 - Then we can also consider clauses or phrases.
 - Example: hotel reviews

Opinion Mining Tasks (cont.)

- At the **feature level** (Example: product reviews, usually you want to know opinions on various features of the product to improve or to compare)
 - *Task 1: Identify and extract object features that have been commented on by an opinion holder*
 - *Task 2: Determine whether the **opinions on the features** are positive, negative or neutral.*
 - *Task 3: Group feature synonyms.*
- **Opinion holders**: identify holders is also useful, e.g., in news articles, etc, but they are usually known in the user generated content, i.e., **authors of the posts.**

Feature-Based Opinion Summary

“I bought an iPhone a few days ago. It was such a nice phone. The touch screen was really cool. The voice quality was clear too. Although the battery life was not long, that is ok for me. However, my mother was mad with me as I did not tell her before I bought the phone. She also thought the phone was too expensive, and wanted me to return it to the shop. ...”

Feature-Based Summary:

Feature1: Touch screen

Positive: 212

- The touch screen was really cool.
- The touch screen was so easy to use and can do amazing things.

...

Negative: 6

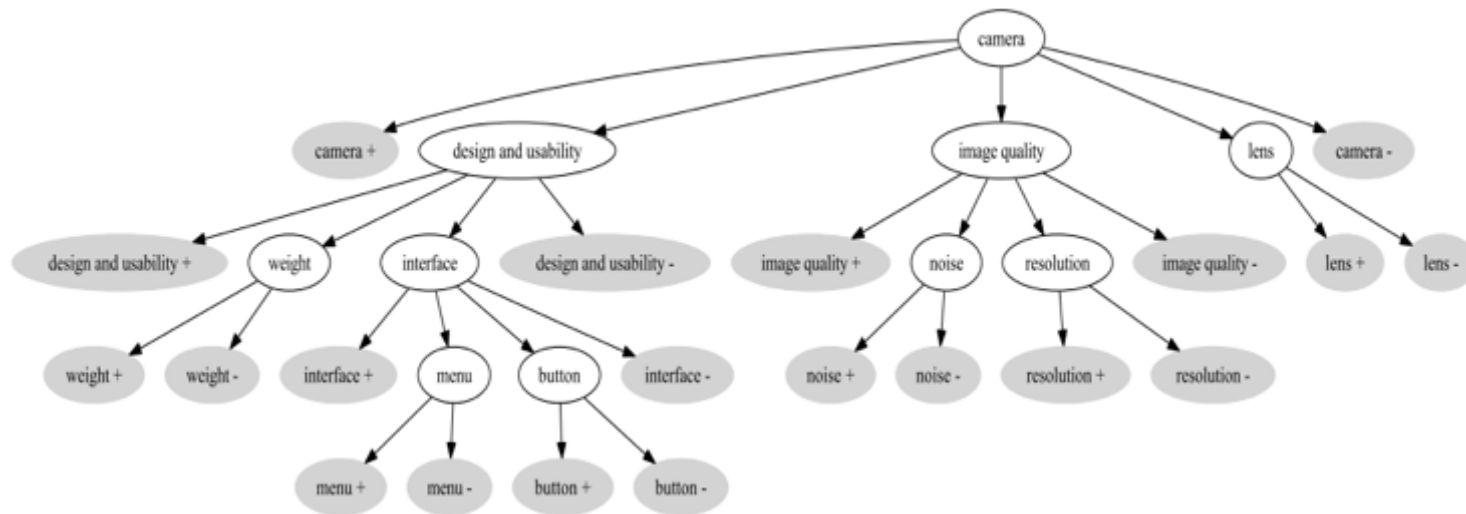
- The screen is easily scratched.
- I have a lot of difficulty in removing finger marks from the touch screen.

...

Feature2: battery life

...

Needs “knowledge” to represent object features



Opinion Analysis: Methods

- Two approaches to the problem:
 1. Machine-Learning (ML) solutions
 2. Lexicon-based solutions
 3. Hybrid solutions
- Each has advantages and disadvantages...

Data Analytics solutions

- Classification
 - Provide an algorithm with lots of **examples**
 - Documents that have been *manually/semi-automatically* annotated with a category
 - *Supervised* learning
 - In our case: e.g., positive/negative reviews (e.g. Tripadvisor)
 - Algorithm extracts “characteristic patterns” for each category and builds a *predictive model*
 - Apply *model* to new text -> get prediction

ML for document classification

- **Bag-of-words** document representation:
document \rightarrow vector (“opinion” words can be considered, or, any word)
 - Example:
 - d_1 = “good.... average... excellent.. good..”
 - d_2 = “okay ..good.. average.. fine..”
 - d_3 = “good... okay.. okay...”
 - Then **Vocabulary** = {“good”, “average”, “excellent”, “fine”, “okay”} and d_1 will be represented as:
 - $d_1 = \{2, 1, 1, 0, 0\}$ if features are frequency-based or
 - $d_1 = \{1, 1, 1, 0, 0\}$ if boolean-based

Bag-of-sentiment-words

I **love** this movie. It's **sweet** but with **satirical** humor. The dialogue is **great** and the adventure scenes are **great fun**...It manages to be **whimsical** and **romantic** while **laughing** at the conventions of the fairy tale genre. I would **recommend** it to just about anyone. I've seen it **several** times as I **love** it so much, and I'm always **happy** to see it **again** whenever I have a friend who hasn't seen it yet.

great	2
love	2
recommend	1
laugh	1
happy	1
.....

Not only words

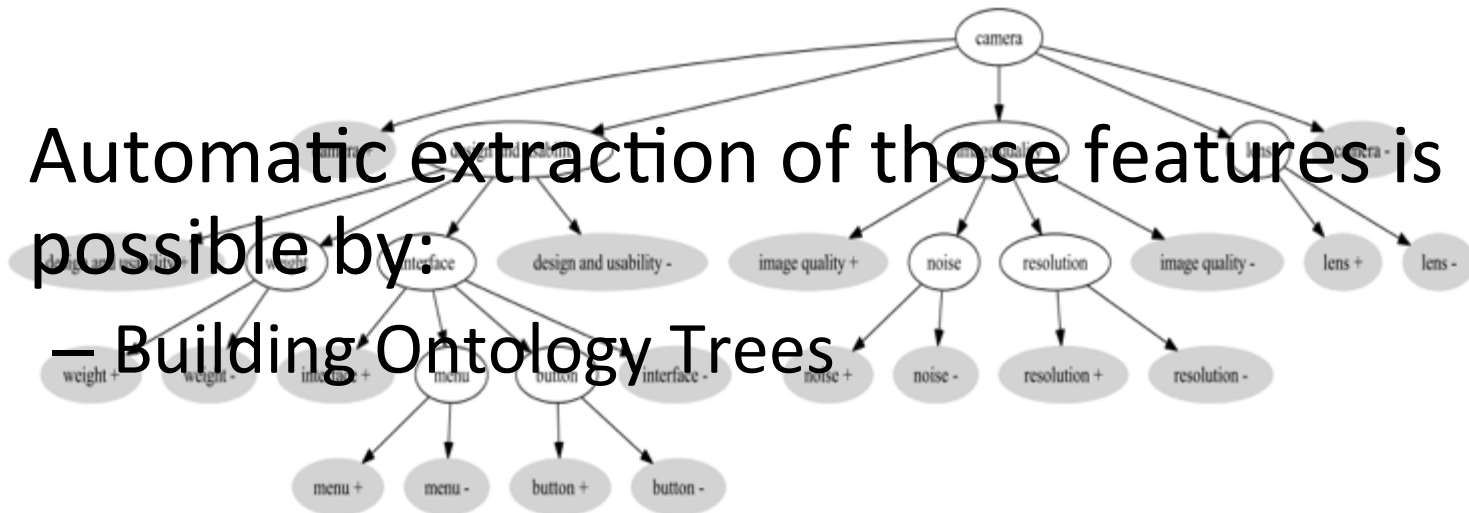
- Typical extensions, focus on *extending/enhancing* the document representation. Instead of/in addition to bag-of-words features, can use:
 - Extra features for *emphasised words* , *special symbols*
 - *wooooow*
 - *exclamations: !!!! ??*
 - *Emoticons* 😊

Feature-based Opinion Analysis

- As discussed, often the *Opinion Object* has different features
 - e.g., camera: lens, quality, weight.
- Often, such an aspect-based analysis is more valuable than a general +/-

- Automatic extraction of those features is possible by:

– Building Ontology Trees



Pros/Cons of the approach

- Advantages:
 - Tend to attain good predictive accuracy
 - Assuming you avoid the typical ML mishaps (e.g., over/underfitting)
- Disadvantages:
 - Need for **training corpus**
 - Solution: automated extraction (e.g., Amazon reviews, Rotten Tomatoes) or crowdsourcing the annotation process (e.g., Mechanical Turk)
 - **Domain sensitivity**
 - Trained models are well-fitted to particular product category (e.g., electronics) **but underperform if applied to other categories** (e.g., tourism)
 - Solution: train a lot of domain-specific models or apply *domain-adaptation* techniques
 - Particularly for Opinion Retrieval, you'll also need to identify the domain of the query!

Example:
"small" is
positive for a
camera,
negative for an
hotel room

Lexicon-based solutions

- Detect/extract the polarity of opinions, based on **affective** dictionaries
- Word-lists where each token is annotated with an ‘emotional’ value
 - e.g., positive/negative words or *words that express anger, fear, happiness, etc.*

Examples of affective dictionaries follow...

- Add **syntactic** and **prose** rules to estimate the overall polarity of text:
 - Negation detection: “the movie **wasn’t** good”
 - Exclamation detection: “great show**!!**”
 - Emoticon detection: “went to the movies 😊”
 - Emphasis detection: “You are go**ooooo**d”
 - Intensifier, diminisher word detection: “**Very** good movie” vs. “good movie”

(Basic) lexicon-based approach

1. Detect emotion in two independent dimensions (numbers are weights of positive/negative opinionated words):
 1. Positive: $D_{\text{pos}}: \{1, 2, \dots, 5\}$
 2. Negative: $D_{\text{neg}}: \{-5, -4, \dots, -1\}$
2. (optional) Predict overall polarity by comparing them :
 - If $D_{\text{pos}} > |D_{\text{neg}}|$ then positive
 - Example: “He is brilliant but boring”
 - Emotion(‘brilliant’)=+3
 - Emotion(‘boring’)=-2

$D_{\text{pos}} = +3, D_{\text{neg}} = -2 \Rightarrow \text{positive}$
3. Negation detection: “He isn’t brilliant and he is boring”
 - *Emotion(NOT ‘brilliant’) = -2*
 - *Decreased by 1 and sign reversed*
4. Exclamation detection: “He is brilliant but boring!!”
 - *Increase weight of emphasized words*
 - *‘boring’=-3*

Pros/Cons of the approach

- Advantages:
 - Can be fairly accurate independent of environment
 - No need for training corpus
 - Can be easily extended to new domains with additional affective words
 - e.g., “amazeballs”
 - More often used in Opinion Retrieval
- Disadvantages:
 - Compared to a well-trained, in-domain ML model they typically underperform
 - Sensitive to affective dictionary coverage

Affective Lexicons

- They have been extensively used in the field either for lexicon-based approaches or in machine-learning solutions
 - Additional features
 - Bootstrapping: unsupervised solutions (see previous)
- Can be created manually, automatically or semi-automatically
- Can be domain-dependent or independent
- A lot of them are already available:
 - Manual
 - LIWC: Linguistic Inquiry and Word Count
 - ANEW: Affective norms for English words
 - Automatic:
 - WordNet-Affect
 - SentiWordNet ...

LIWC: Linguistic Inquiry and Word Count

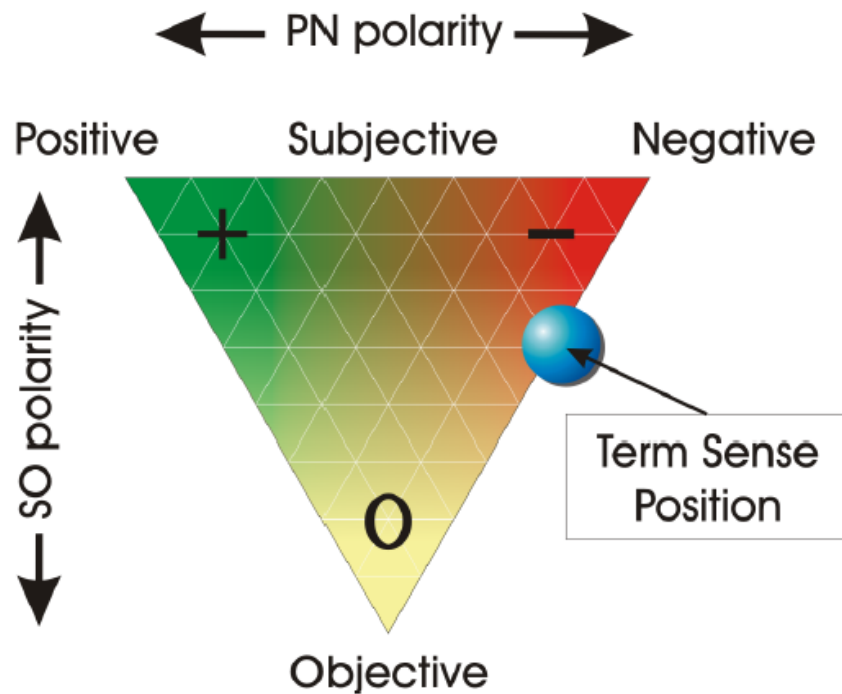
125 Affect						126 Posemo			127 Negemo			
abandon*	damn*	fume*	kind*	privleg*	supporting	accept	freed*	partie*	abandon*	enrag*	maddening	snob*
abuse*	danger*	fuming	kiss*	prize*	supportive*	accepta*	freeing	party*	abuse*	envie*	madder	sob
abusi*	daring	fun	laidback	problem*	supports	accepted	freely	passion*	abusi*	envious	maddest	sobbed
accept	darlin*	funn*	lame*	profit*	suprem*	accepting	freeness	peace*	ache*	envy*	maniac*	sobbing
accepta*	daze*	furious*	laugh*	promis*	sure*	accepts	freer	perfect*	aching	evil*	masochis*	sobs
accepted	dear*	fury	lazie*	protest	surpris*	active*	frees*	play	advers*	excruciat*	melanchol*	solemn*
accepting	decay*	geek*	lazy	protested	suspicio*	admir*	friend*	played	afraid	exhaust*	mess	sorrow*
accepts	defeat*	genero*	liabilit*	protesting	sweet	ador*	fun	playful*	aggravat*	fail*	messy	sorry
ache*	defect*	gentle	liar*	proud*	sweetheart*	advantag*	funn*	playing	aggress*	fake	miser*	spite*
aching	defenc*	gentler	libert*	puk*	sweetie*	adventur*	genero*	plays	agitat*	fatal*	miss	stammer*
active*	defens*	gentlest	lied	punish*	sweetly	affection*	gentle	pleasant*	agoniz*	fatigu*	missed	stank
admir*	definite	gently	lies	radian*	sweetness*	agree	gentler	please*	agony	fault*	misses	startl*
ador*	definitely	gigg*	like	rage*	sweets	agreeab*	gentlest	pleasing	alarm*	fear	missing	steal*
advantag*	degrad*	giver*	likeab*	raging	talent*	agreed	gently	pleasur*	alone	feared	mistak*	stench*
adventur*	delectabl*	giving	liked	rancid*	tantrum*	agreeing	gigg*	popular*	anger*	fearful*	mock	stink*
advers*	delicate*	glad	likes	rape*	tears	agreement*	giver*	positiv*	angr*	fearing	mocked	strain*
affection*	delicious*	gladly	liking	raping	teas*	agrees	giving	prais*	anguish*	fears	mockers*	strange
afraid	deligh*	glamor*	livel*	rapist*	tehe	alright*	glad	precious*	annoy*	feroc*	mocking	stress*
aggravat*	depress*	glamour*	LMAO	readiness	temper	amaz*	gladly	prettie*	antagoni*	feud*	mocks	struggl*
aggress*	depriv*	gloom*	LOL	ready	tempers	amor*	glamor*	pretty	anxi*	fiery	molest*	stubborn*
agitat*	despair*	glori*	lone*	reassur*	tender*	amus*	glamour*	pride	apath*	fight*	mooch*	stunk
agoniz*	desperat*	glory	longing*	rebel*	tense*	aok	glori*	privleg*	appall*	fired	moodi*	stunned
agony	despis*	goddam*	lose	reek*	tensing	appreciat*	glory	prize*	apprehens*	flunk*	moody	stuns
agree	destroy*	good	loser*	regret*	tension*	assur*	good	profit*	argh*	foe*	moron*	stupid*
agreeab*	destruct*	goodness	loses	reject*	terribl*	attachment*	goodness	promis*	argu*	fool*	mourn*	stutter*
agreed	determina*	gorgeous*	losing	relax*	terrific*	attract*	gorgeous*	proud*	arrogan*	forbid*	murder*	submissive*
agreeing	determined	gossip*	loss*	relief	terrified	award*	grace	radian*	asham*	fought	nag*	suck
agreement*	devastat*	grace	lost	reliev*	terrifies	awesome	graced	readiness	assault*	frantic*	nast*	sucked
agrees	devil*	graced	lous*	reluctan*	terrify	beaut*	graceful*	ready	asshole*	freak*	needy	sucker*
alarm*	devot*	graceful*	love	remorse*	terrifying	beloved	graces	reassur*	attack*	fright*	neglect*	sucks
alone	difficult*	graces	loved	repress*	terror*	benefic*	graci*	relax*	aversi*	frustrat*	nerd*	sucky
alright*	digni*	graci*	lovely	resent*	thank	benefit	grand	relief	avoid*	fuck	nervous*	suffer

ANEW: Affective norms for English words

Description	Word No.	Valence Mean(SD)	Arousal Mean(SD)	Dominance Mean (SD)	Word Frequency
abduction	621	2.76 (2.06)	5.53 (2.43)	3.49 (2.38)	1
abortion	622	3.50 (2.30)	5.39 (2.80)	4.59 (2.54)	6
absurd	623	4.26 (1.82)	4.36 (2.20)	4.73 (1.72)	17
abundance	624	6.59 (2.01)	5.51 (2.63)	5.80 (2.16)	13
abuse	1	1.80 (1.23)	6.83 (2.70)	3.69 (2.94)	18
acceptance	625	7.98 (1.42)	5.40 (2.70)	6.64 (1.91)	49
accident	2	2.05 (1.19)	6.26 (2.87)	3.76 (2.22)	33
ace	626	6.88 (1.93)	5.50 (2.66)	6.39 (2.31)	15
ache	627	2.46 (1.52)	5.00 (2.45)	3.54 (1.73)	4
achievement	3	7.89 (1.38)	5.53 (2.81)	6.56 (2.35)	65
activate	4	5.46 (0.98)	4.86 (2.56)	5.43 (1.84)	2
addict	581	2.48 (2.08)	5.66 (2.26)	3.72 (2.54)	1
addicted	628	2.51 (1.42)	4.81 (2.46)	3.46 (2.23)	3
admired	5	7.74 (1.84)	6.11 (2.36)	7.53 (1.94)	17
adorable	6	7.81 (1.24)	5.12 (2.71)	5.74 (2.48)	3
adult	546	6.49 (1.50)	4.76 (1.95)	5.75 (2.21)	25
advantage	629	6.95 (1.85)	4.76 (2.18)	6.36 (2.23)	73
adventure	630	7.60 (1.50)	6.98 (2.15)	6.46 (1.67)	14
affection	7	8.39 (0.86)	6.21 (2.75)	6.08 (2.22)	18
afraid	8	2.00 (1.28)	6.67 (2.54)	3.98 (2.63)	57

sentiwordnet.isti.cnr.it/

SentiWordNet



Opinion-Mining Tools



Pulse of the Nation: U.S. Mood Throughout the Day inferred from Twitter

Less Happy  More Happy

<http://www.ccs.neu.edu/home/amislove/twittermood>

<http://www.ccs.neu.edu/home/amislove/twittermood/>

Twitter investor sentiment



Investor Sentiment

@Tweet_Sentiment

Helping investor navigate the Twitter Sentiment. AKA, using twitter to predict the predict a Bull or Bear market. Invest real-time using this twitter feed!

chocolatechipapps.com

Iscritto a febbraio 2011


Twitta a Investor Sentiment

TWEET 24 FOLLOWING 472 FOLLOWER 137 Altro ▾


Tweet Tweet e risposte

Ritwittato da Investor Sentiment
 **Scott Eddy** @MrScottEddy · 4 feb 2011
Congress Grills Facebook On Plans To Share User Addresses, Cell Numbers <http://huff.to/eFC0Cu>


← ↻ 1 ★ ⋮

Ritwittato da Investor Sentiment
 **Yahoo Finance** @YahooFinance · 4 feb 2011
Stocks up ahead of US jobs data, Egypt woes loom <http://yhoo.it/hvDrCB> #Futures

← ↻ 5 ★ ⋮

Ritwittato da Investor Sentiment
 **QualityStocks** @QualityStocks · 4 feb 2011
January jobs report forecast to show modest gains <http://ow.ly/3Q0kN> ~ <http://disclaim.it/f/8ewa>

← ↻ 2 ★ 1 ⋮

 **Investor Sentiment** @Tweet_Sentiment · 4 feb 2011
dvolatility: RT @BreakingNews: Suspect in custody after gunman briefly hijacks Greyhound bus in N.C.; no one hurt -... <http://ff.im/xlcwh>

← ↻ ★ 1 ⋮



TRACKING OPINIONS ON TWITTER

twitrratr

SEARCH

Discover what people are really saying on Twitter. With Twitrratr you can distinguish negative from positive tweets surrounding a brand, product, person or topic.

TERM

st ives

POSITIVE TWEETS

70

NEUTRAL TWEETS

384

NEGATIVE TWEETS

11

TOTAL TWEETS

465

15.05% POSITIVE



i really want to love st. ives apricot scrub, but it irritates my skin soo much :([\(view\)](#)



rt @kesiahosking: sunshine was smiling at you annie =)) rt @anniegreenwood st ives harbour basking in november sunshine <http://flic.kr/p/8tk2sq> [\(view\)](#)



sunshine was smiling at you annie =)) rt @anniegreenwood st ives harbour basking in november sunshine <http://flic.kr/p/8tk2sq> [\(view\)](#)



looking at st ives (uk:siv). great stats, but printing? [\(view\)](#)

82.58% NEUTRAL



@oldergirlbeauty GURL, I was all about the Aqua Net & the St. Ives liquid hairspray in the purple bottle. Where's my banana clip? [\(view\)](#)



RT @inscriptions: Loved the final episode of Junior Masterchef! Alexwill be at St Ives Village Sat 11th to show us a thing or two! [\(view\)](#)



Loved the final episode of Junior Masterchef! Alex from top12 is coming to St Ives Village Sat 11th to show us a thing or two about cooking! [\(view\)](#)

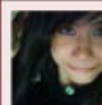


A Town On Canvas Called St Ives <http://ping.fm/onNWl> [\(view\)](#)

2.37% NEGATIVE



st. ives apricot scrub is bad for your face. you may not notice it but it scratches up your face and its bad... <http://bit.ly/dttmci> [\(view\)](#)



st ives face scrub receive negative comments. lots of it o.o [\(view\)](#)



@fandomonymous not sure how bad your acne is, but st. ives green tea cleanser works well on my skin. really cleans out my pores. [\(view\)](#)



sco prem: goal st ives city 2 towerhill blues 0 lucas k (43) [\(view\)](#)



sco prem: goal st ives city 1

tweetfeel



Bonanza

Search

Try some Twitter trends: [Romo](#) [Bonanza](#)



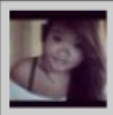
4



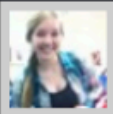
3



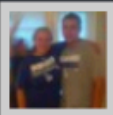
57%



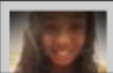
RT @BuildYourLoveUp: RT @BuildYourLoveUp: I wish @itsimreeeee went to the same school as me. I miss my best friend, and almost everyone at [Bonanza](#) sucks. ;/



shout out to the helicopter circling our school this morning with a spot light.. I love [Bonanza](#) <http://t.co/j2EDX0cS>



RT @ADReamGONe: RT @ADReamGONe: Man, I love [Bonanza](#) Imfao.



RT @ADReamGONe: RT @ADReamGONe: Man, I love [Bonanza](#) Imfao.

Sentiment140

 Tweet < 353

 Like < 140

 +1 < 74

English ▾

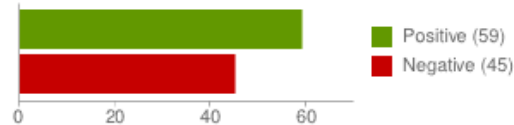
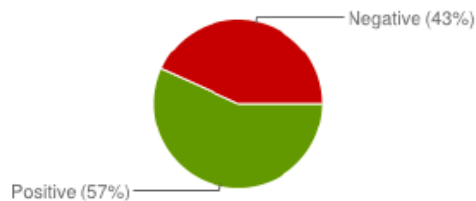
Search

[Save this search](#)

Sentiment analysis for microsoft

Sentiment by Percent

Sentiment by Count



Tweets about: microsoft

Isaydumb: @Youporn, in my humble opinion you have nothing to do on the @Xbox Live. What the fuck is @Microsoft doing?!

Posted 46 seconds ago

Megan Maracle: I hate this class. #Microsoft #die

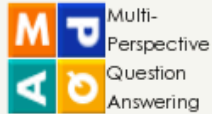
Posted 2 minutes ago

dilwortha: @carasmith10 oh okay, you'll have to explain when i see you as i dont understand this disk haha. is it for microsoft project do you

Posted 5 minutes ago

jlebrech: @rsslldnphy it happens to be microsoft this time, but a superset is the next best thing from a compiled bytecode, as valid JS is also

Opinion Finder



Main
MPQA Home

Corpora
News, debates, etc.

Lexicons
Subj. clues, etc.

Annotation
GATE, MPQA scheme

OpinionFinder
Subjectivity detector

OpinionFinder

Version 1.x

[Version 1.5](#)

[Version 1.4](#)

[Sample Annotations](#)

[Version 2.x](#)

OpinionFinder 1.x Release Page

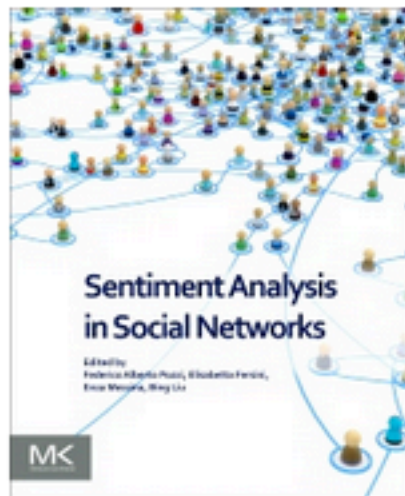
OpinionFinder 1.x Available versions

OpinionFinder 1.x relies on many external software packages (e.g. SUNDANCE, SCOL, BoosTexter) which are neither built nor supported by our group. Since OpinionFinder was originally released in 2005, there are some compatibility issues with versions of various software and packages. We have reports that these problems sometimes result in an exhausting and even unsuccessful installation process of OpinionFinder. Since many of the people involved in the original development have graduated and left the group, we do not currently have the resources to address these compatibility issues concerning the required external software packages. Although we do not have the resources to bring OpinionFinder 1.x fully up-to-date, we are currently working on a new version of OpinionFinder. OpinionFinder 2 is being written in Java and will be platform-independent.

LICENSE AGREEMENT

Version 1.5

- [README - OpinionFinder 1.5](#)
- [Download OpinionFinder 1.5](#)



Sentiment Analysis in Social Networks

1st Edition

[View on ScienceDirect](#) ↗



Authors: Federico Alberto Pozzi, Elisabetta Fersini, Enza Messina, Bing Liu

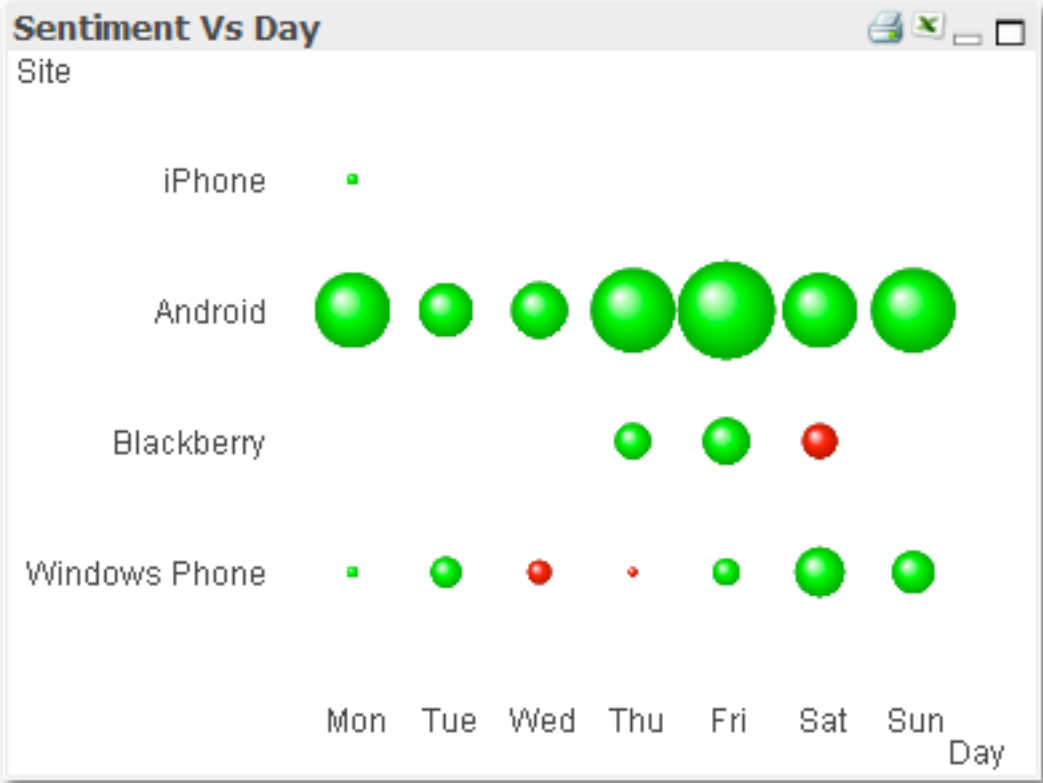
eBook ISBN: 9780128044384

Paperback ISBN: 9780128044124

Imprint: Morgan Kaufmann

Published Date: 15th September 2016

Page Count: 284



Apple Watch - Words

Top Words by Volume

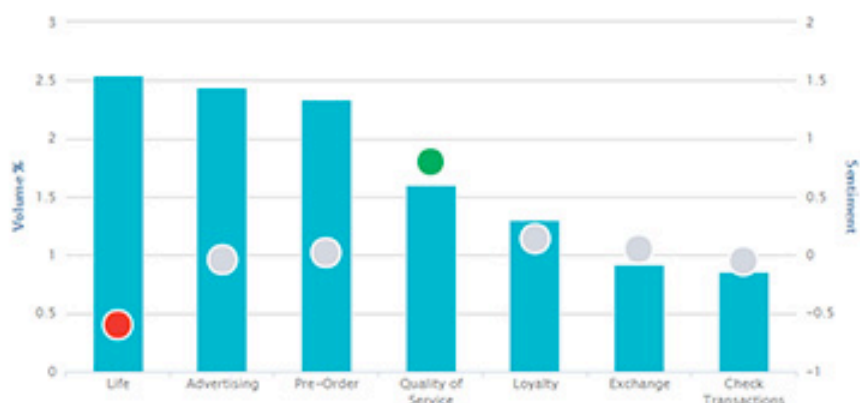
As of 10/10/14



Volume and Sentiment - Key Topics

Top Topics by Volume %

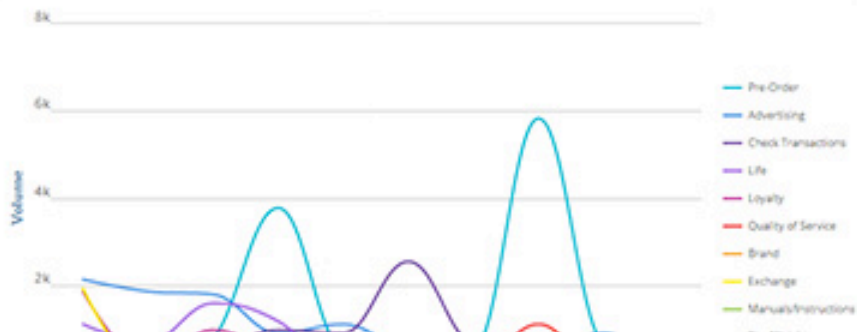
As of 10/8/14



Topic Trending - Last 10 Days

9/9/14 to 9/18/14

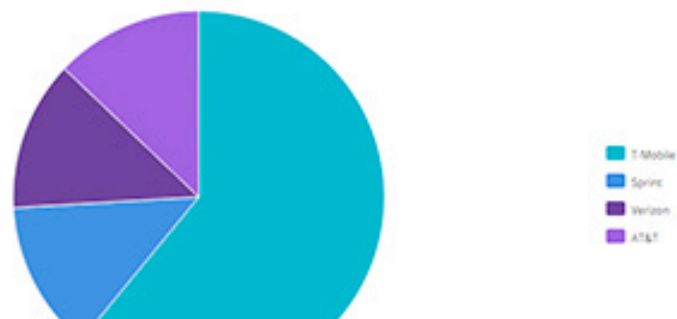
As of 10/12/14



Topics Report - Model: Telephone Retailers

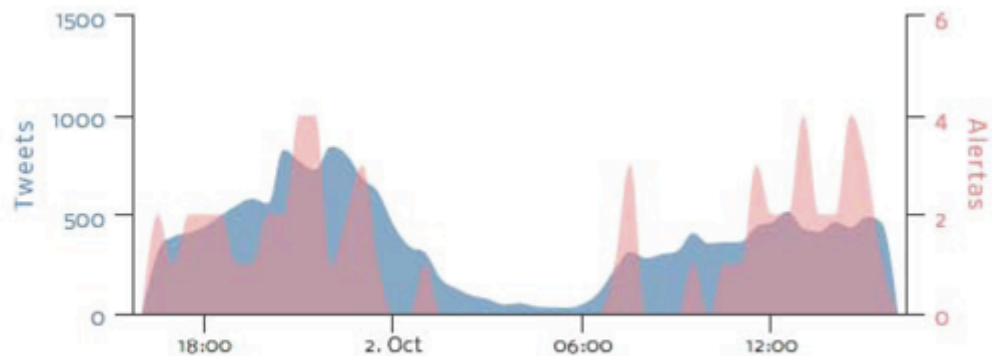
Top Topics by Volume

As of 10/12/14



17903
TWEETS

62
ALERTAS



Alertas +

actos delictivos - narcotráfico 14

accidente - accidente en carretera

8

★ Entidades +

AP-7 #PontdeMolins

Usuarios +

InfoEmerg bomberscat mossoscat
transfap7

Conceptos +

Hypertext Transfer Protocol accidente
camión clase cocaína crimen funeral
inundación retuiteo terrorista
violencia de género víctima

Hashtags +

PontdeMolins SEM

Temas +

policía y justicia 20

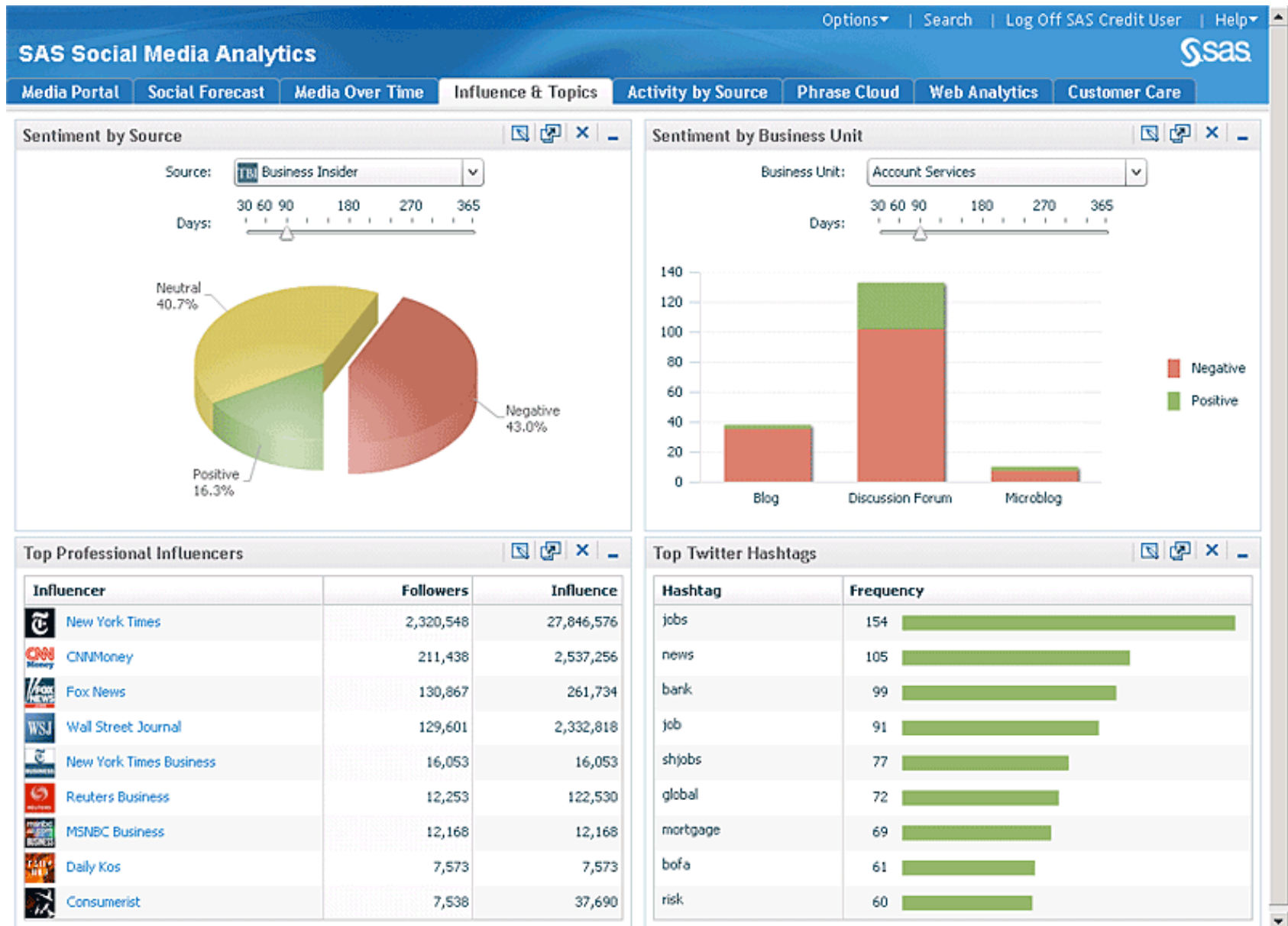
catástrofes y accidentes 18

PLATFORMS FOR SENTIMENT ANALYSIS

<https://www.ibm.com/support/knowledgecenter/SSJHE9>



http://www.sas.com/en_us/industry/media.html



http://www.opinioneq.com/

OpinionEQ GPS Features COMP Actions ▾

Studies & Reports << **Study Overview** Study Settings

Studies Reports by Type

- GPS Studies
 - Demo GPS Study
 - Standard Reports
 - Custom Reports
 - Test Garmin
 - Standard Reports
 - Washers
 - Washer
 - Standard Reports
 - Washer (copy)
 - BLOOM Studies
 - Blackberrys
 - Standard Reports
 - Miller Beer
 - Standard Reports
 - Will Studies
 - Sony TVs (copy)
 - Standard Reports
 - COMP Studies
 - GPS Features COMP**
 - Standard Reports
 - Custom Reports

Overall Sentiment

Sentiment	Percentage
Strongly Pos.	~25%
Positive	~55%
Negative	~15%
Strongly Neg.	~5%

Overall Trend

Month	Opinions	% Pos. Sentiment
Feb	~400	~74%
Mar	~450	~78%
Apr	~500	~73%
May	~550	~74%
Jun	~600	~75%
Jul	~650	~70%
Aug	~650	~71%
Sep	~650	~75%
Oct	~600	~71%
Nov	~550	~72%
Dec	~650	~74%
Jan	~600	~75%
Feb	~350	~76%

Opinions

the case seems very sturdy : in fact all the pieces appear to be made of quality plastic

(Some of the less expensive Garmins do not come with the USB cable , which is a problem because

to be fair , the TomTom is an older unit , but it was a much more expensive unit than the Garmin 285

the device also comes the MSN Direct (9 months free) which I found to be of very limited usefulness

the difference is that the 285WT comes with the MSN receiver (with Some free months , then a month

If you drive as much as I do , especially around unfamiliar roads , you know how important it is to know

There have been times when the prices have turned up wrong , and it only shows the price of Regular

the main advance in GPS systems Over the last three years has been price reduction

Id say the 9 free months of MSN Direct is better to have than not , but basically this unit should be priced

Great buy , great price , great service

Page 1 of 1085 | Displaying opinions 1 - 10 of 10845

Frequent Words

Top Features | Top Brands | Top Products

am amazon best better bluetooth bought buy cheaper cost device

download easy even excellent expensive features firmware first

free garmin good gps great masellan map maps model

money new nice navi one price product quality really

roads screen size software time tom tomtom traffic unit

update updates value works worth

<http://www.netvibes.com>



Semantria (free demo available)

Text Analytics Demo

This demo shows some of the text analytics features available via our [services](#).
Sign up for a [30 day trial](#) of Semantria for Excel and API to explore and customize the full [output](#).

Let's start by analyzing a single document:

This document is: **neutral (+0.010)**

English

No Industry Pack

<https://www.nytimes.com/2017/10/12/us/politics/trump-obamacare-> [Go](#)

Highlight: Phrases Themes Entities

WASHINGTON — President Trump will scrap subsidies to **health** insurance companies that help pay out-of-pocket costs of **low-income people**, the White House said **late** Thursday. His plans were disclosed hours after the president ordered **potentially** sweeping changes in the nation's insurance system, including sales of cheaper policies with fewer **benefits** and fewer protections for consumers.

The twin hits to the Affordable Care Act could unravel President Barack Obama's signature domestic **achievement**, sending insurance **premiums soaring** and insurance companies fleeing from the **health** law's online marketplaces. After Republicans **failed** to **repeal** the **health** law in Congress, Mr. Trump appears determined to dismantle it on his own.

Without the subsidies, insurance markets could **quickly** unravel. Insurers have said they will need much higher **premiums** and may pull out of the insurance exchanges created under the Affordable Care Act if the subsidies were **cut off**. Known as cost-sharing reduction payments, the subsidies were expected to total \$9 billion in the coming year and nearly \$100 billion in the coming decade.

"The government cannot **lawfully** make the cost-sharing reduction payments," the White House said in a statement.

Current Character Count: 11884 / 16384

[Clear](#) [Start Analysis](#)

sabotage easier
small businesses **very happy**
cancer mistake **failed**
spiteful new products
punish troubled **ill**
individual employees
professional **disastrous**
make it **comprehensive**
nightmare happy praised

Scroll down for full report



Other platforms

- More on <https://blog.hootsuite.com/social-media-sentiment-analysis-tools/>
- <http://barnraisersllc.com/2017/01/best-tools-sentiment-analysis-free-fee/>