TU Graz Guest Lecture

Max Pellert

2023-11-07

Max Pellert

Max Pellert   About   Courses & Talks   Materials

Max Pellert

Assistant Professor
University of Mannheim

Download CV as PDF

Bio

Max Pellert has a background in cognitive science and economics (University of Vienna, Austria and University of Ljubljana, Slovenia). He was a doctoral researcher affiliated to Complexity Science Hub Vienna and Medical University of Vienna in the WWTF research group “Emotional Well-Being in the Digital Society” lead by David Garcia (now University of Konstanz). After receiving his PhD, he gained industry experience as Assistant Researcher at Sony CSL Rome. Currently, he works at the Chair for Data Science in the Economic and Social Sciences at University of Mannheim (Markus Strohmaier) as junior faculty (assistant professor). His research focuses on analyzing the digital traces of individual and collective emotional behavior and affective expression on social media. He is broadly interested in the social sciences and uses traditional and novel computational methods from domains such as Natural Language Processing to study emotion dynamics, belief updating, collective emotions and other interesting phenomena.

https://mpellert.at

Max Pellert

Interdisciplinary background: BSc Economics (and studies in Psychology), MSc Cognitive Science and a PhD in Computational Social Science

All of the degrees are from Vienna (University of Vienna and Medical University of Vienna), semester abroad in Ljubljana, Slovenia

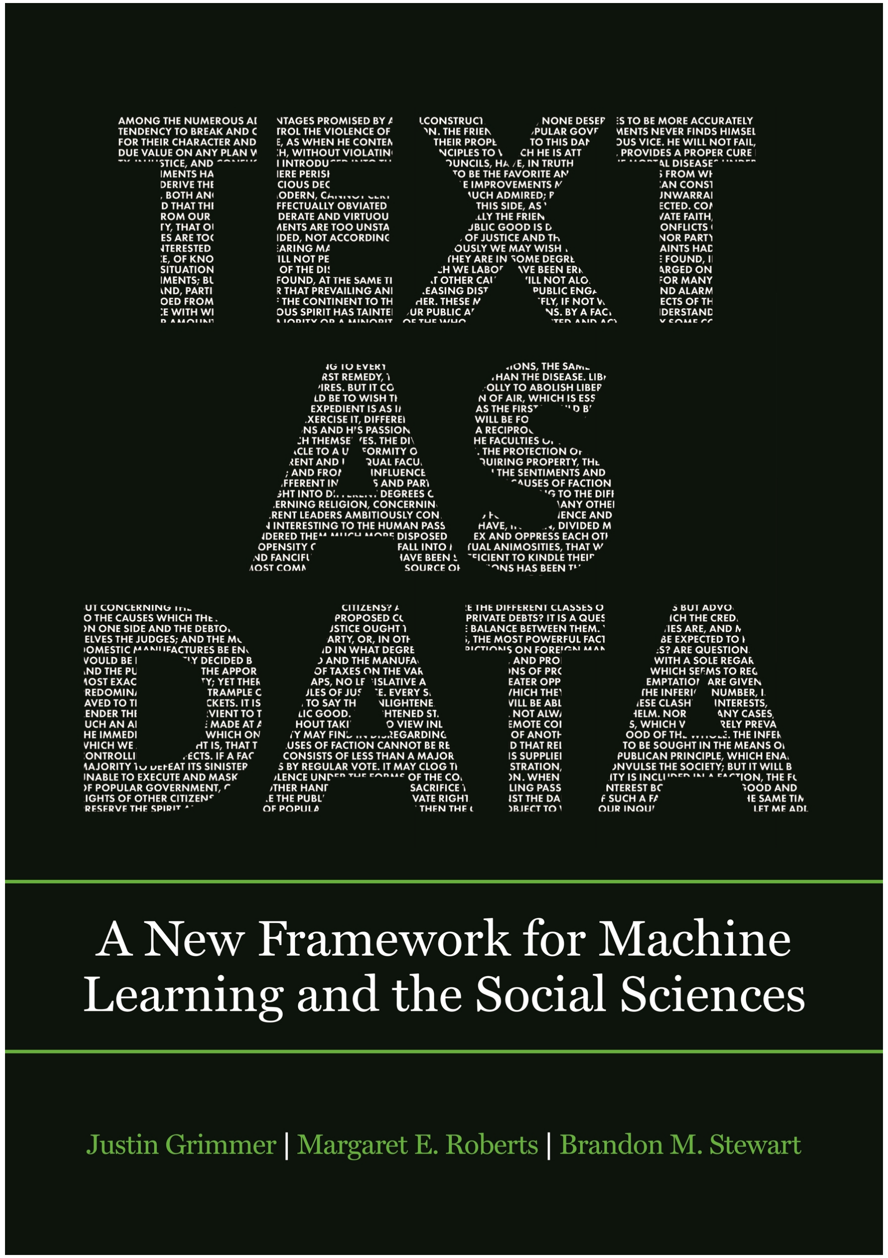
Before coming to the University of Mannheim as an assistant professor, I worked at Sony Computer Science Laboratories in Rome, Italy

Chair for Data-Science in the Economic and Social Sciences, Business School **University of Mannheim**

Max Pellert

Research interests

- Computational Social Science
- Digital traces
- Affective expression in text
- Natural Language Processing
- Collective emotions
- Belief updating
- **Psychometrics of AI**


TEXT AS DATA
A New Framework for Machine Learning and the Social Sciences
Justin Grimmer | Margaret E. Roberts | Brandon M. Stewart

“Text as Data”

- The book can help to find out what methods are available and how they can (and have been) used to tackle research questions

- Starts with meta-theoretical considerations and gives you some kind of roadmap on how to use text as data to tackle scientific questions

“Text as Data”

- Builds up sophisticated machinery by going from simple to advanced in a very concise, efficient way (providing lots of pointers to additional materials)

- Can serve as a work of reference to look up certain methods that you might need and get inspiration on how to use them (for example different clustering techniques are covered in one of the chapters)

What are emotions?

> Emotions as **core affect**: Short-lived psychological states that consume the individual's energy and strongly influence cognition and behavior, for example expression.

Emotional or affective behavior of an individual takes place at various timescales:

What are emotions?

- Reflex reactions: fast physiological responses
- Core affect: relax quickly and are triggered by a stimulus
- Mood: slow-changing and constant emotional state
- Personality traits are lifelong behavior patterns, some about emotions

Computational Affective Science

> **Affective Science** is the (interdisciplinary) scientific study of emotions.

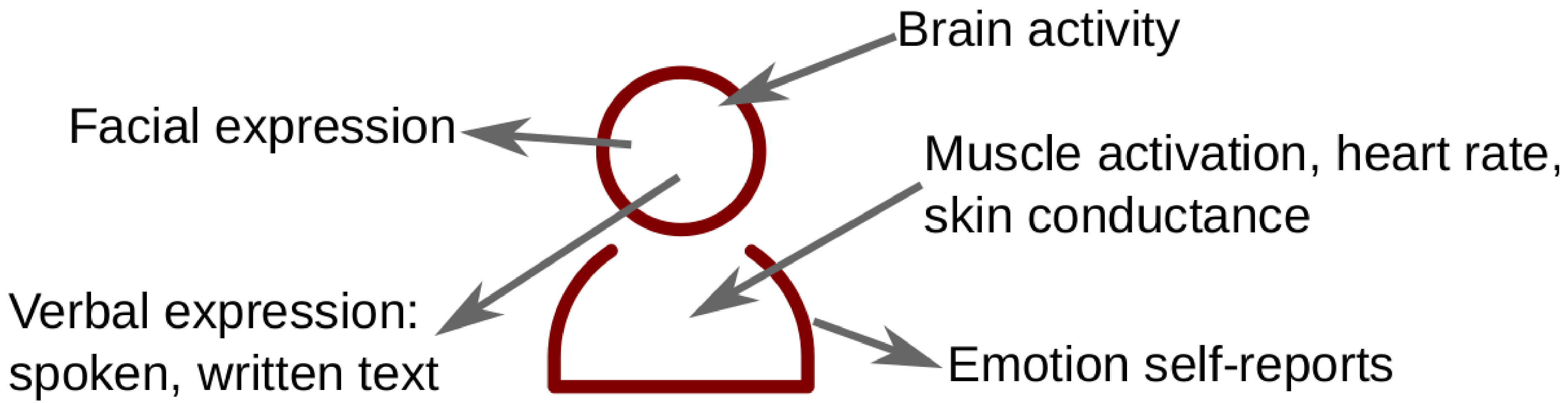
> **Computational Affective Science** applies methods from Computer Science and Data Science to Affective Science. Some examples are:

Computational Affective Science

- **Affective Computing:** Development of systems that detect, process, and elicit emotion
- **Cyberpsychology of Affect:** Understanding the interplay between emotions and ICT
- **Emotion Recognition:** Identification of human emotion using any kind of modality: text, voice, facial expression, physiological signals (skin conductance, muscle activity, EEG, fMRI), etc
- **Sentiment Analysis:** Detection of subjective states from (textual) data, including emotion

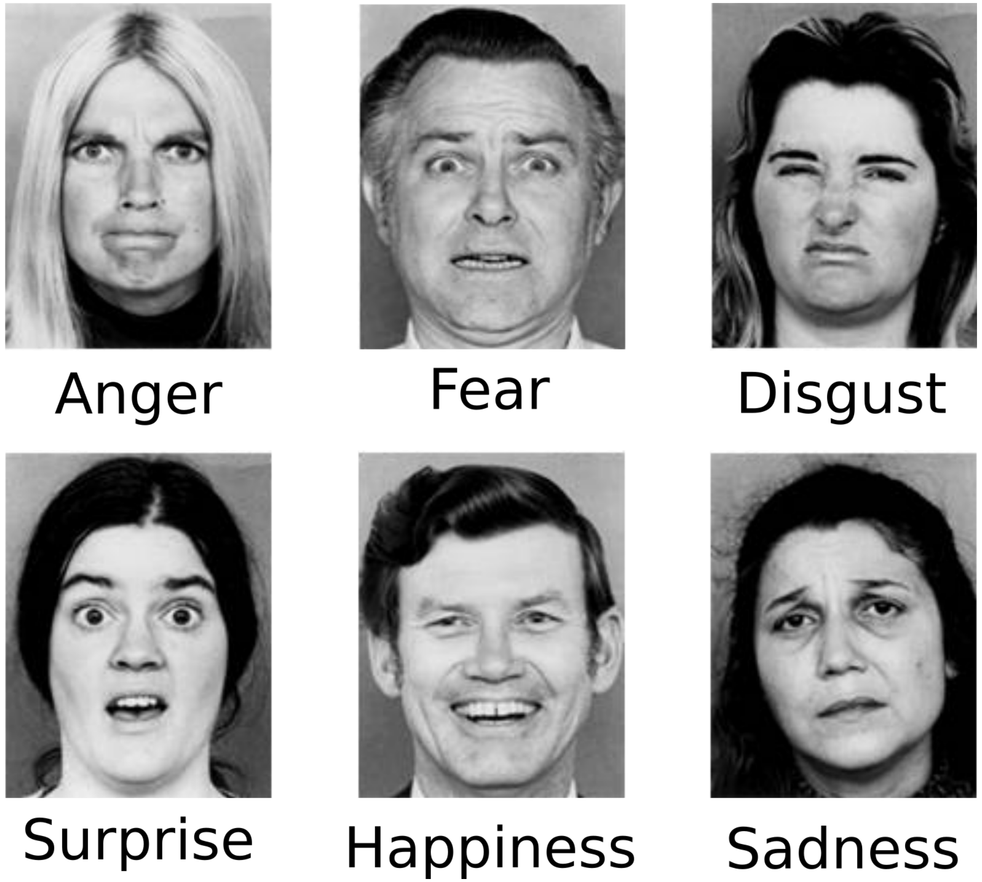
Measuring emotions

Emotions can be measured through various signals and observable behaviors:



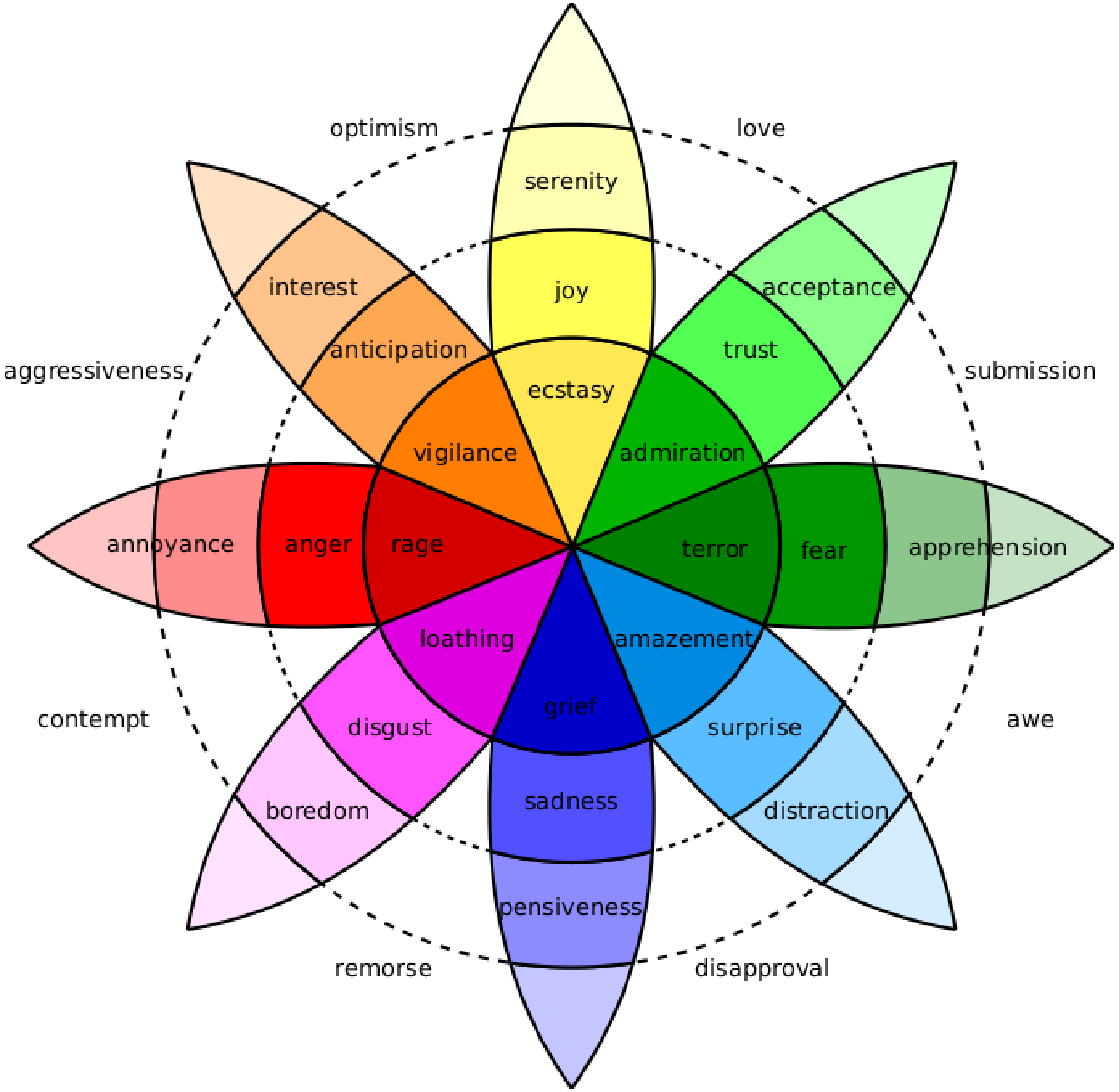
In the following, we are going to cover four models of how to capture emotions in quantitative research. Some approaches are better for some modes or signals (e.g. text, facial expression) than others.

Ekman's basic emotions model

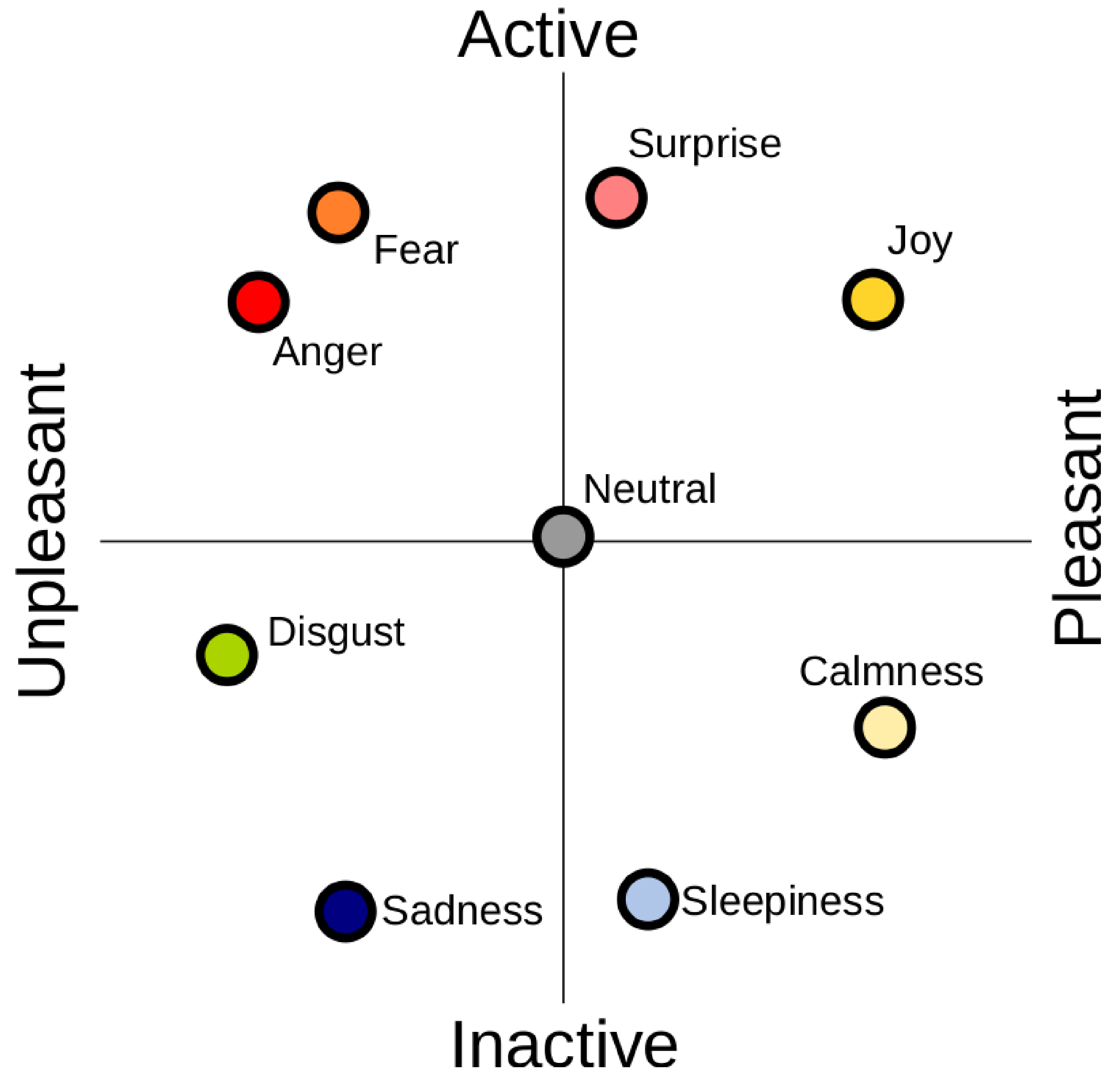


Developed by **Paul Ekman** to classify facial expression of emotions.

Plutchik's wheel of emotions



The circumplex model of affect



Dimensions in the circumplex model

> **Valence:** the degree of pleasure experienced in an emotion

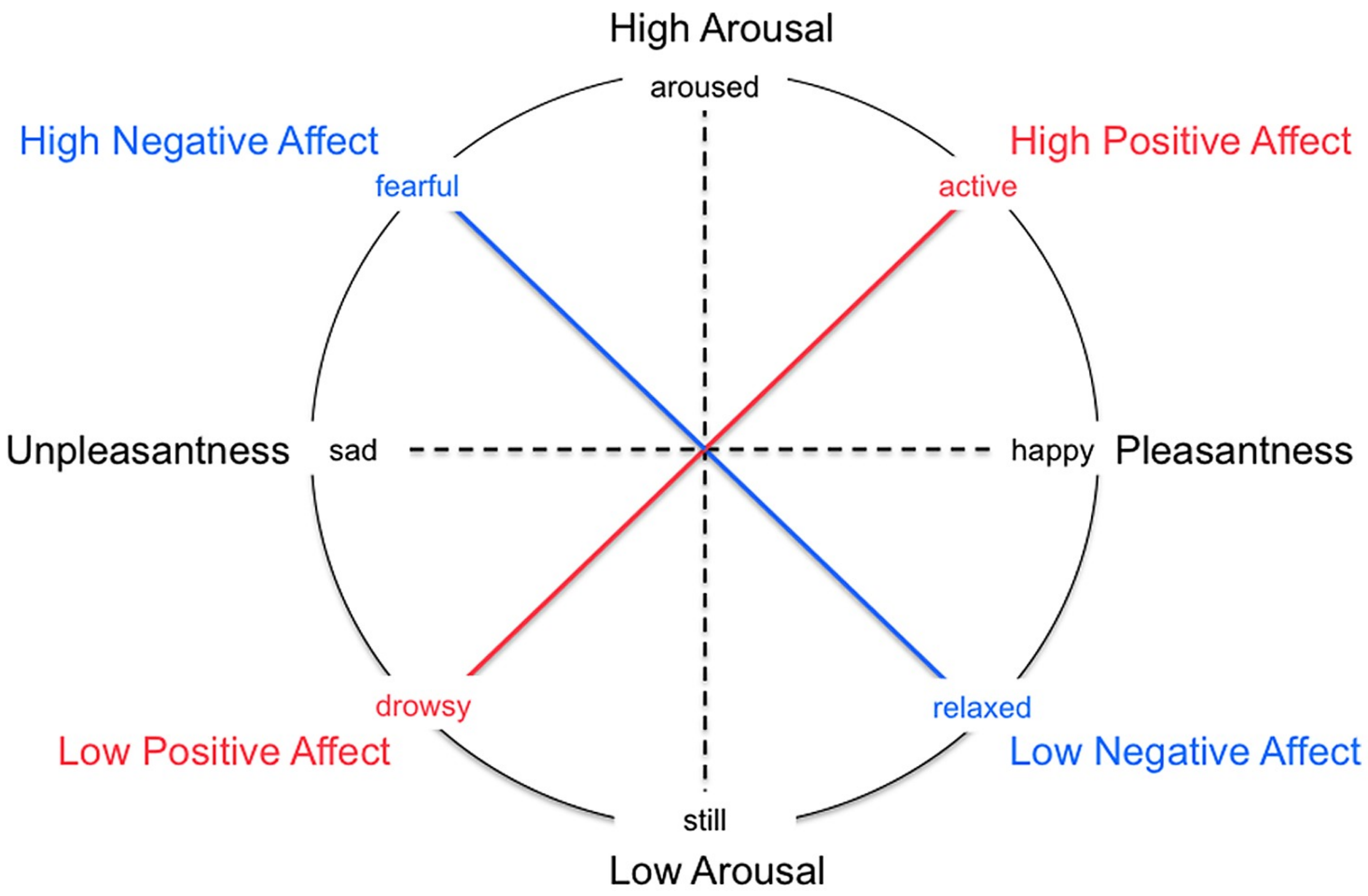
- Explains the most variance from positive/pleasant to negative/unpleasant
- It can be measured physiologically through muscle activity associated with smiling and frowning
- It is the most common dimension of emotions included in text analysis

Dimensions in the circumplex model

> **Arousal:** the level of activity associated with an emotion

- Explains less variance than valence but is informative to differentiate emotions
- It can be measured with skin conductance and heart rate sensors
- Not so common in text analysis but it can for example be estimated from voice tone

Positive And Negative Affect Schedule



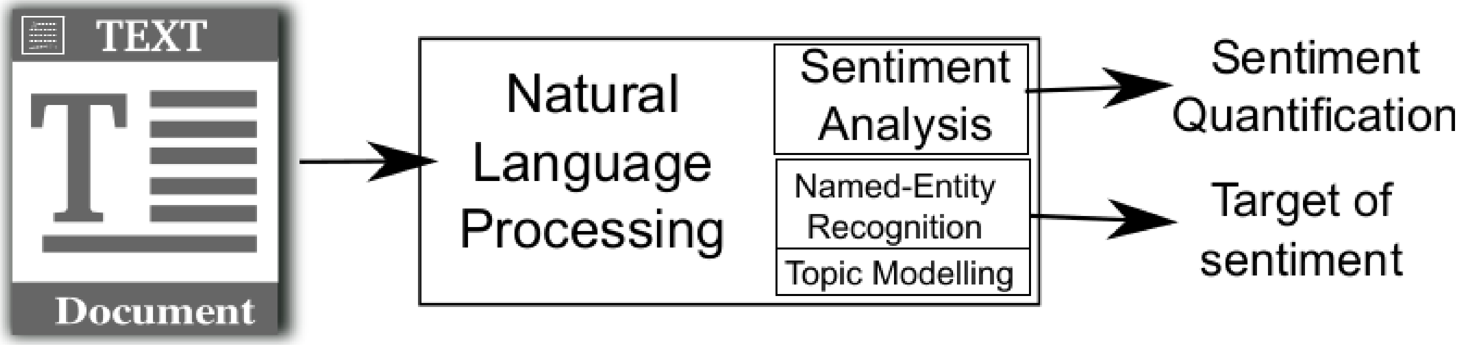
What is Sentiment Analysis?

> **Sentiment Analysis:** Computerized quantification of subjective states from text

- Examples of subjective states: Emotions, feelings, attitudes, opinions…
- Often vaguely defined and roughly equivalent to the dimension of valence in Russell's model
- Sentiment quantification can have various formats
  - Polarity of the text: positive, negative, or neutral
  - Numeric scores of positive and negative content
  - Labels of emotions in text: e.g. Joy, Anger, Sadness, Disgust, Surprise…

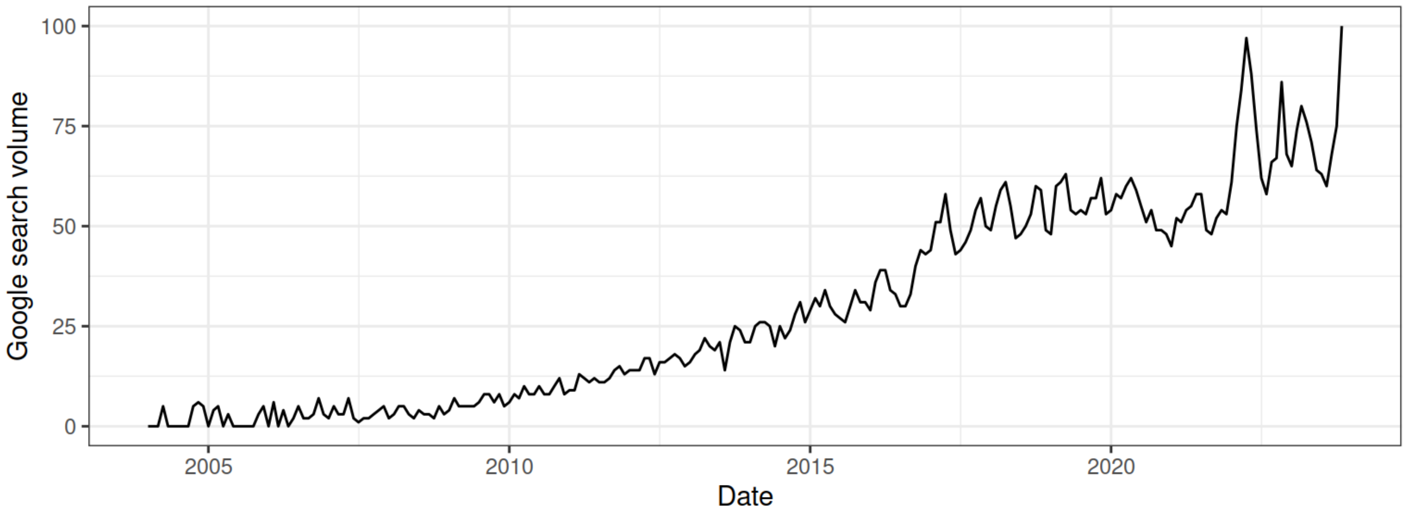
What is Sentiment Analysis?

Sentiment analysis is a subfield of **Natural Language Processing**. It can be combined with other tools like Named-Entity Recognition or Topic Modelling to contextualize the sentiment, for example finding its origin or targets. Here we focus on how to quantify sentiment from text, especially in social media and other kinds of digital traces.



What is Sentiment Analysis?

There has been a scientific boom in sentiment analysis with several workshops, journal issues, and books devoted to the topic. Every year there are hundreds of research papers on the topic. You can see this rise in the Google Trends volume for the term “sentiment analysis”:



Supervised and unsupervised sentiment analysis

Unsupervised sentiment analysis:

- Uses expert knowledge (e.g. from psychologists) to quantify emotions
- Expert knowledge is encoded as a set of rules or a lexicon (dictionary) of words
- Pros: Simple implementation, large coverage and recall
- Cons: Hard to customize for a particular context, low precision, expert bias

Supervised and unsupervised sentiment analysis

Supervised sentiment analysis:

- Uses a set of annotated texts to fit a model
- Annotations can come from readers or the authors of texts
- Pros: Automatic calibration, high precision
- Cons: Lower recall and coverage, need very large training datasets

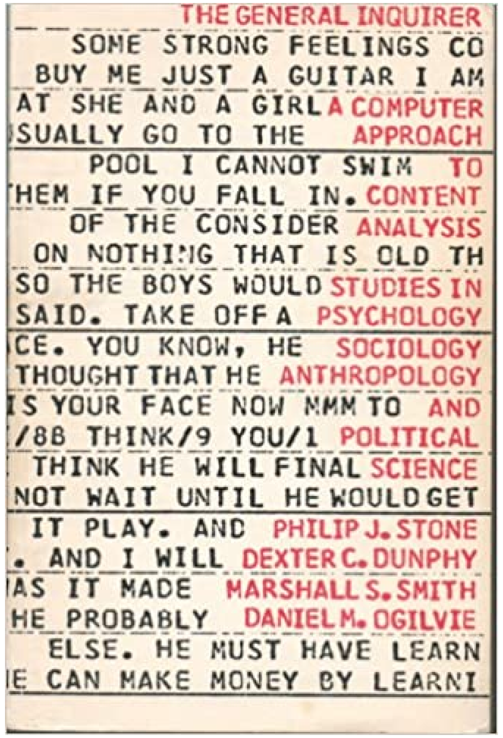
Supervised and unsupervised sentiment analysis

Both approaches can be combined in what is called semi-supervised or ensemble methods. Some of these approaches mix supervised and unsupervised models in one classifier.

Evaluation and generalizability are key arguments when choosing a sentiment analysis method. In this topic we are going to cover various approaches to unsupervised sentiment analysis with examples of methods and software you can use.

The General Inquirer

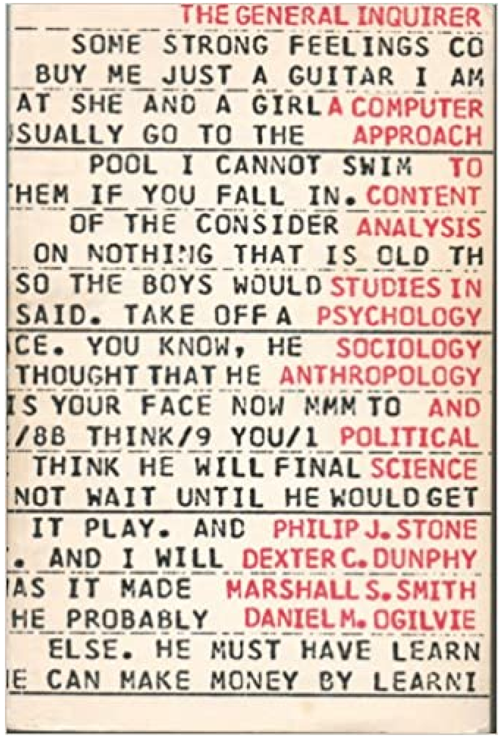
The pioneer work of Philip Stone in 1966 proposed to process text with a computer to detect the use of words of various categories.



This set the basis for **dictionary methods** in unsupervised sentiment analysis, which are based on counting the number of appearances of the words of a list in a text. The original version of the General Inquirer contained many word classes including parts of speech, topics, as well as terms for emotions and evaluative language.

The General Inquirer

The original dictionaries of the General Inquirer were merged with other later dictionaries and an updated version was released in the 1990s. You can access the lists of positive words and of negative words of this version, which served as input for later methods like SentiStrength.



The SentimentAnalysis R package contains the General Inquirer (GI) dictionary and methods to match words in text.

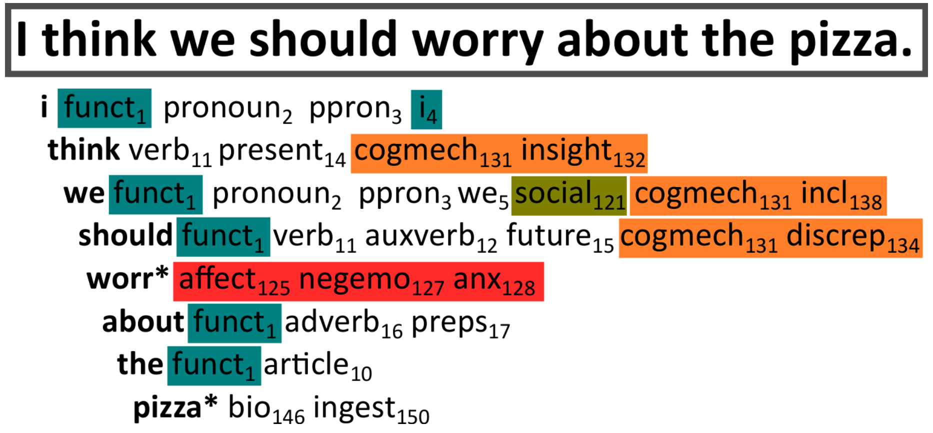
Linguistic Inquiry and Word Count (LIWC)

LIWC (pronounced "Luke") was developed as a click-and-run software by James Pennebaker in 2001.

Inspired by the General Inquirer, it contains a set of word lists that are matched against words in the text to compute frequencies for each list. The word lists of LIWC were designed to cover both linguistic classes and to capture psychological processes such as cognitive processes, social processes, and emotions.

Word lists for LIWC are produced by groups of experts that compare their individual word lists and expand them with synonyms. There have been three versions of LIWC in English (2001, 2007 and 2015) and dictionaries have been generated with the same method for several languages including German, Spanish, French, Arabic, and Chinese.

Here you can see an example of how LIWC words on a text:



LIWC first tokenizes the text, i.e. it identifies words by looking for separations like whitespaces and punctuation. Then LIWC iterates over each token (word) and checks if it matches any word list in the dictionary. These matches can be “hard” matches for the same exact character string, or “soft” matches with Kleene stems that are prefixes of a word. These are entries in the dictionary that end with a star symbol (“*”). You see this in the example for the entry “worr*” that matches “worry” and for “pizza*” that matches “pizza”.

Linguistic Inquiry and Word Count (LIWC)

In the example you could see that words can belong to several word lists, for example the entry for “worr*” is in the “affect” list, in the “negemo” list, and in the “anxiety” list. After running these matchings, LIWC produces a list of frequency measures as the percentage of words in the whole text that are matched against each word list. In the example above, there are 12.5% words of the “negemo” list and 0% words of the “posemo” list.

The 2015 version of LIWC includes netspeak terms such as “WTF” or “LOL” and emoticons like “:)”, LIWC is a very popular tool due to the ease to use it, for example it offers a way to visualize which words are matched. It is very important to look at these matches to understand LIWC emotion word frequencies, as you can learn when we talk about emotions in pagers after 9/11.

SentiStrength

Mike Thelwall developed in 2010 the SentiStrength method: a sentiment analysis method designed to quantify positive and negative sentiment from short, informal social media text.

Test- Download - Java Version - Non-English - Buy! - About

SentiStrength

Automatic sentiment analysis of up to 16,000 social web texts per second with up to human level accuracy for English - other languages available or easily added.

SentiStrength estimates the *strength* of positive and negative sentiment in *short texts*, even for informal language. It has human-level accuracy for short social web texts in English, except political texts.

SentiStrength processess text in three steps:

1. Text preprocessing: correcting misspellings, vowel repetitions, translating emoticons and idioms, etc
2. Match words from scored list of words in the scale [-5,+5]
3. Apply modifiers (negation, amplification, de-amplification). These modifiers change the polarity of words and their strength. The final scores are an aggregate of these polarities.

SentiStrength takes two sources of expert input: a word list with sentiment scores and a list of modifier rules including terms for negation, amplification, etc. SentiStrength outputs two scores: a positive score [+1,+5] and negative score [-1,-5]. Ths design matches the PANAS scales, you can learn more about them in the Measuring Emotions topic.

Sentistrength has been adapted and validated for various languages including Spanish, German, and Russian. It is distributed as a Java executable with available code and can be run from the command line with text files as input.

VADER (Valence Aware Dictionary and sEntiment Reasoner)

VADER is a tool very similar to SentiStrength, tailored to detect sentiment on Twitter by C.J. Hutto and Eric Gilbert in 2014. It applies the same three steps as SentiStrenght:

VADER tutorial

1. Text preprocessing
2. Word matching from a lexicon of positive/negative scored words
3. Application of modifiers to the scores based on language rules

VADER (Valence Aware Dictionary and sEntiment Reasoner)

VADER's name suggests it is the "dark version" of LIWC ("Luke"). As the authors of VADER say: *"VADER distinguishes itself from LIWC in that it is more sensitive to sentiment expressions in social media contexts."*

VADER was implemented in Python and distributed as an open source package on Github and as part of the NLTK python library for NLP. Its performance was validated against annotated tweets, correlating the scores given by tweet readers with the output of VADER. VADER can be run in R with the package vader.

The digital traces of pagers

Back in the 90s, pagers were a common form of mobile communication. To send a message to a pager, you could call a special phone number, say your message, and the text of the message would appear in the screen of the pager.

Emotions in pagers after 9/11

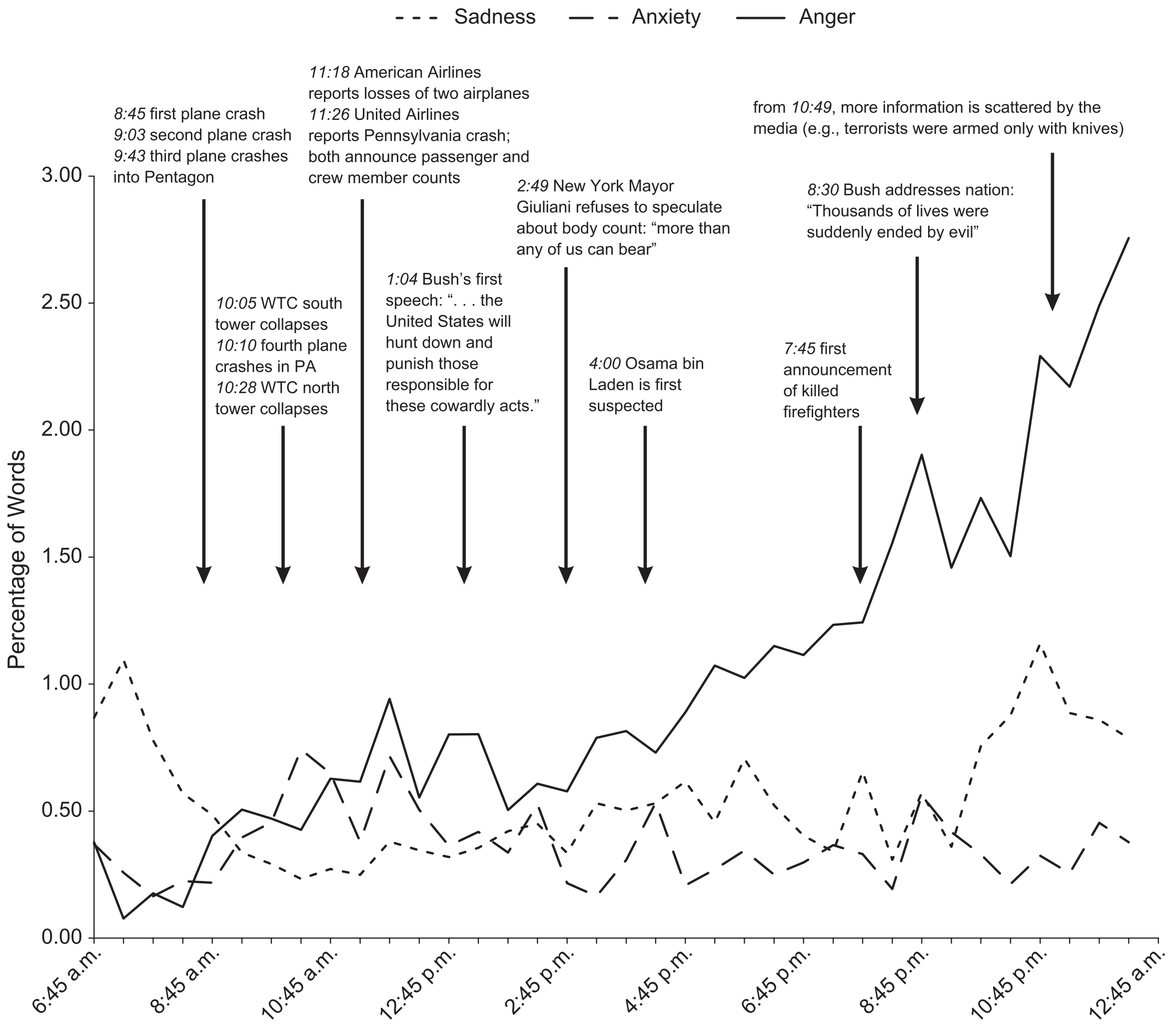


Fig. 1. The timeline of sadness, anxiety, and anger on September 11 as expressed in messages sent to text pagers. Each data point represents the mean percentage of words related to the specific negative emotion, averaged across 30 min. The time slots start at 6:45 a.m. to 7:14 a.m. on September 11, 2001, and end at 12:15 a.m. to 12:44 a.m. on September 12, 2001. Exact times and brief descriptions of the most important events of September 11 are included above the timelines. WTC = World Trade Center

Not so angry americans

More than a third of anger words appeared in messages like these:

2001-09-12 02:25:12 Arch [0987275] C ALPHA
s0191: 09/11 13:18:30 Reboot NT machine gblnetnt05 in cabinet 311R at 13/1CMP:CRITICAL:Sep 11 13:18:30

2001-09-12 02:25:14 Arch [0987275] C ALPHA
s0191: 09/11 13:19:18 Reboot NT machine gblnetnt07 in cabinet 311R at 13/1CMP:CRITICAL:Sep 11 13:19:18

2001-09-12 02:25:16 Arch [0951146] C ALPHA
TX-013 Caddo - No answer.

2001-09-12 02:25:16 Arch [0987275] C ALPHA
s0191: 09/11 13:27:06 Reboot NT machine gblnetnt06 in cabinet 311R at 13/1CMP:CRITICAL:Sep 11 13:27:06

2001-09-12 02:25:16 Arch [1657188] B ALPHA
NCAGBRWS03:Device Inaccessible

2001-09-12 02:25:17 Skytel [007569153] A ST
NUM 413-397-1947-143823

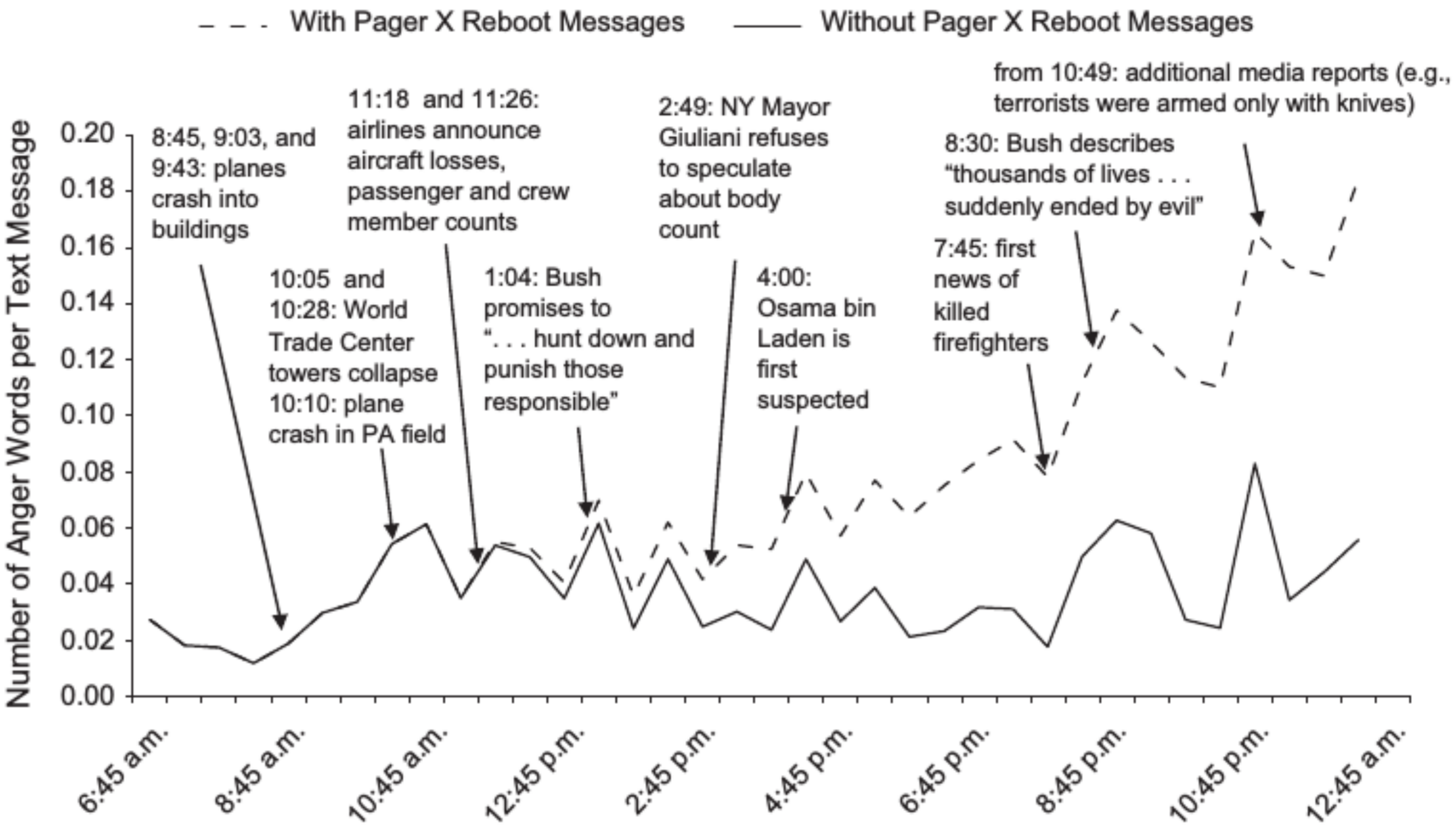
2001-09-12 02:25:17 Skytel [005192078] D ALPHA
thanks! -michael b.

2001-09-12 02:25:17 Skytel [004693762] A ALPHA
CALL MY CELL PHONE WHEN YOU CAN. I LOST MY PAGER. I LOVE YOU. 31.

"Reboot NT machine [name] in cabinet [name] at [location]:CRITICAL:[date and time]."

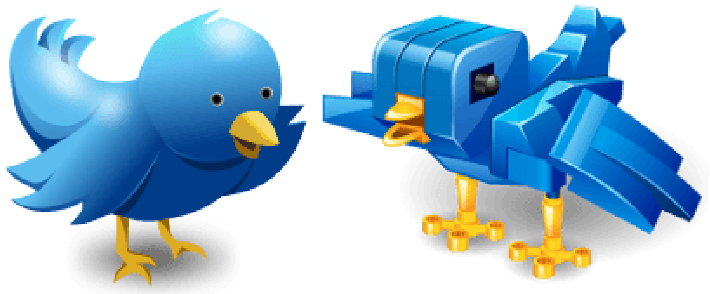
The word "critical" is contained in the anger word list of LIWC!

Anger timeline without REBOOT messages



The issue of machine-generated traces

Not all digital traces are generated by humans, a large volume of data is generated by machines.

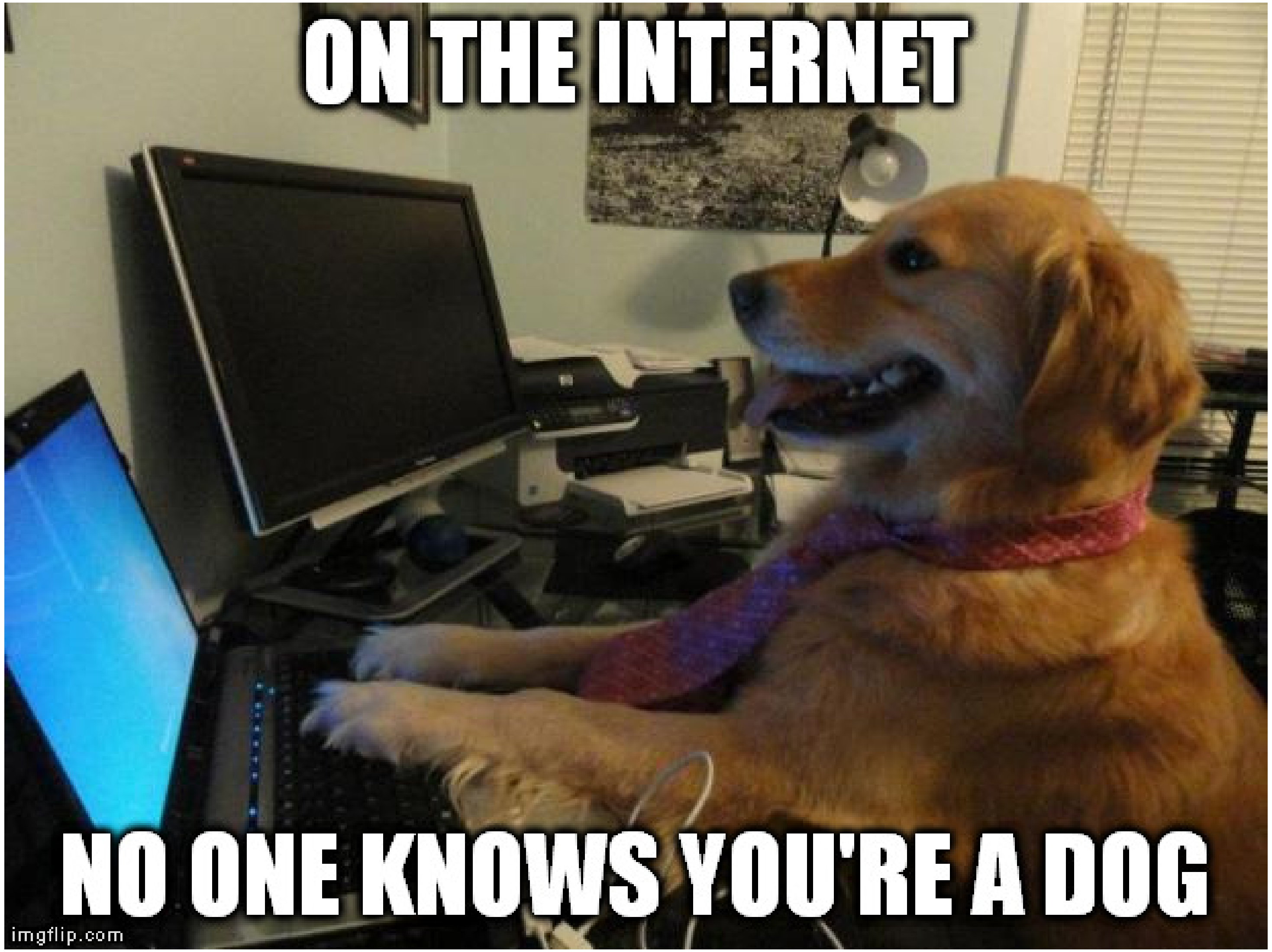


During the summer of 2018, Twitter made a big bot cleanse, but independent estimates before reported that between 9% and 15% of Twitter accounts were likely to be bots.

The issue of machine-generated traces

One of the most widely used methods to detect bots on Twitter is Botometer, which is in constant development by the OSoMe lab at Indiana University. Even if you clean bots this way from your data, you should always take a good look at your text. You can make word clouds, word shift graphs, or just browse through it to see if you notice anomalous patterns. To sum up:

> Take home message: Do not just analyze text, also look at it!

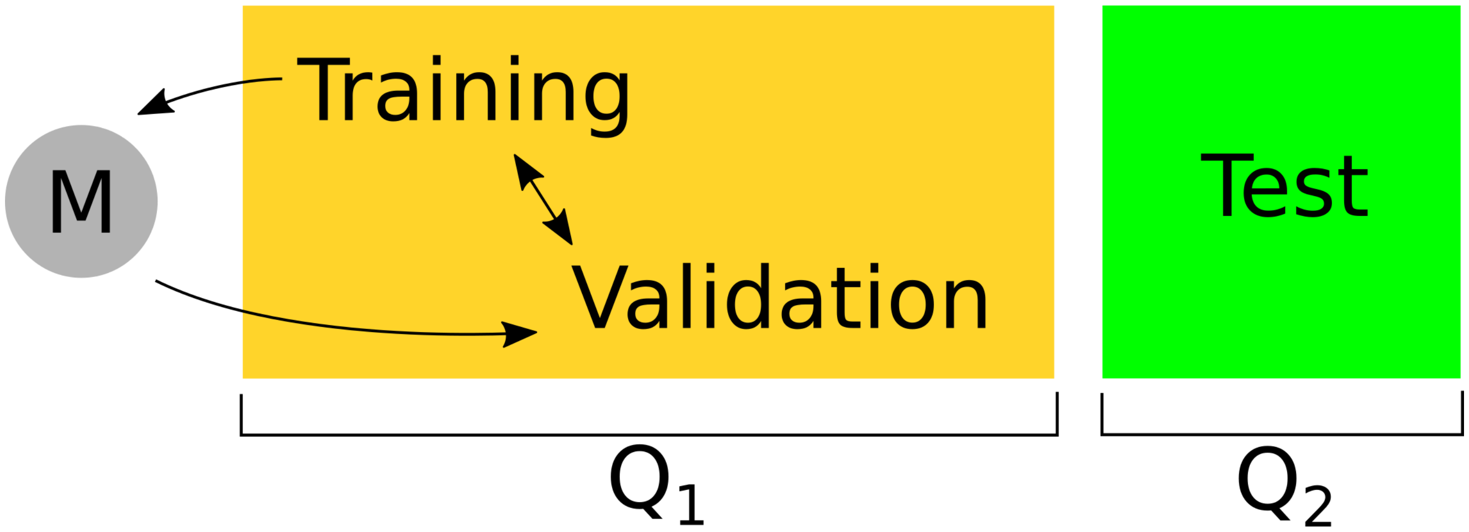

ON THE INTERNET
NO ONE KNOWS YOU'RE A DOG
imgflip.com

Summary so far

- **Measuring emotions**
  - Modalities of emotion measurement
  - Emotions as core affect: the circumplex model
  - Other models: Ekman's model, Plutchik's wheel, PANAS
- **Unsupervised sentiment analysis**
  - Method in reference to other methods in NLP
  - Dictionary-based methods: the General Inquirer and LIWC
  - Rule-based methods: Sentistrength and VADER
- **The case of emotions in pagers after 9/11**
  - An application of LIWC to a hacked dataset of digital communication
  - Example of systematic error due to one word
  - Call to inspect text and consider bots

Supervised approaches

How to train your model



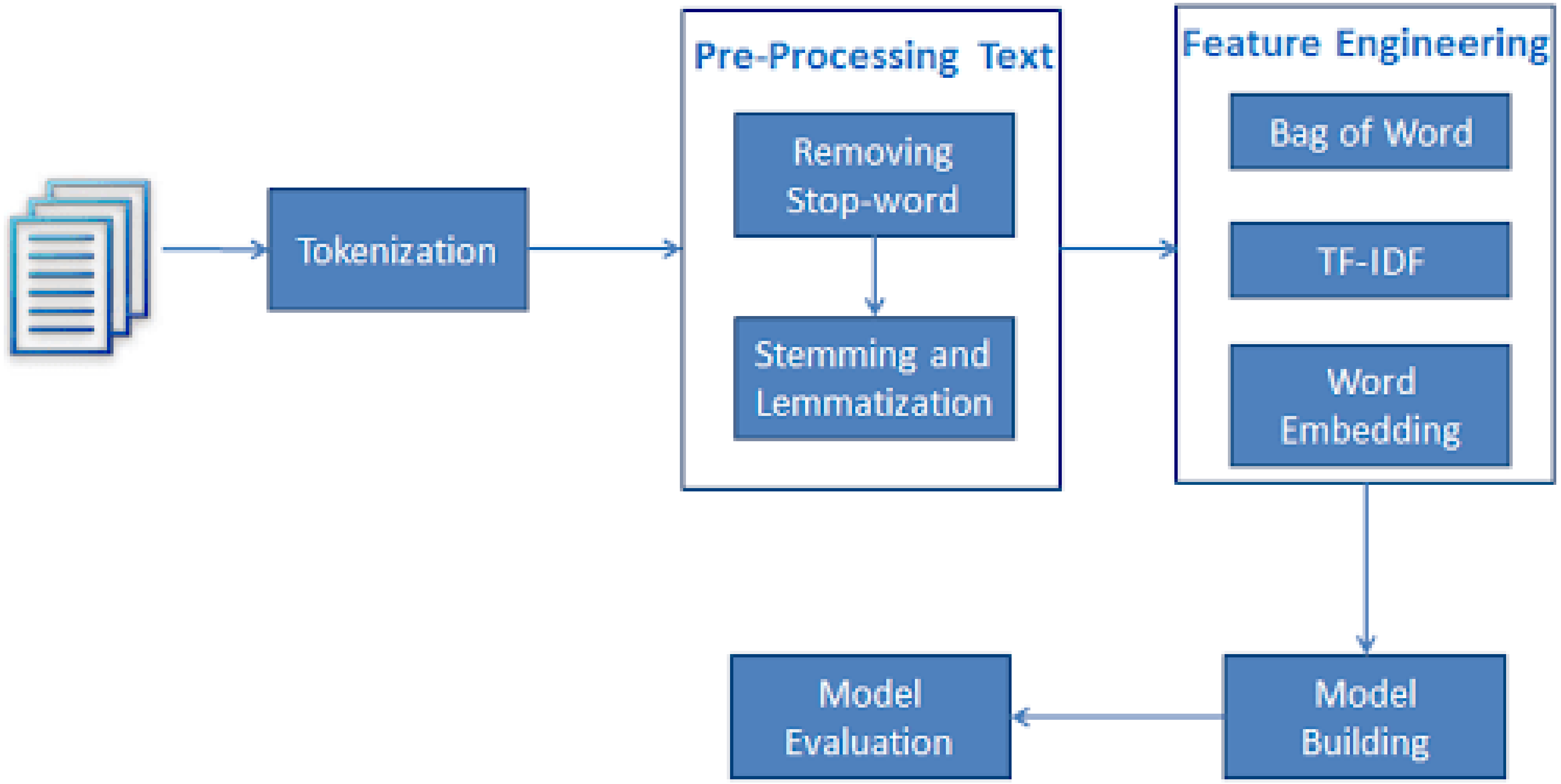
1. **Training:** Texts with annotations of sentiment are used to fit the model
2. **Validation:** A second set of annotated texts not used for training are used to evaluate the quality of the model: Q_1
3. **Testing:** One last evaluation run of the fitted model with a leave-out sample of annotated texts. None of these texts should have been part of the validation or training steps. Testing quality Q_2 measures predictive quality of the model.

Gold Standards and Ground Truths

pos	neg	text
0	0	I wana see the vid Kyan
0	1	I cant feel my feet.. that cant be good..
1	0	10 absolutely jaw dropping concept car designs http://ow.ly/15OnX
0	0	Phil Collins- You Can't Hurry Love

- Supervised sentiment analysis needs a set of labeled texts to learn from.
- Labels can come from the author of the text or from reading raters
- The above table is an example from a real dataset with two annotations: a positivity score and a negativity score
- Other ground truths might have numeric scores in a scale or text labels for basic emotional states.

Text preprocessing



Pre-processing from Text Analytics for Beginners using NLTK, Navlani, 2019

What model to use?

Common approaches are:

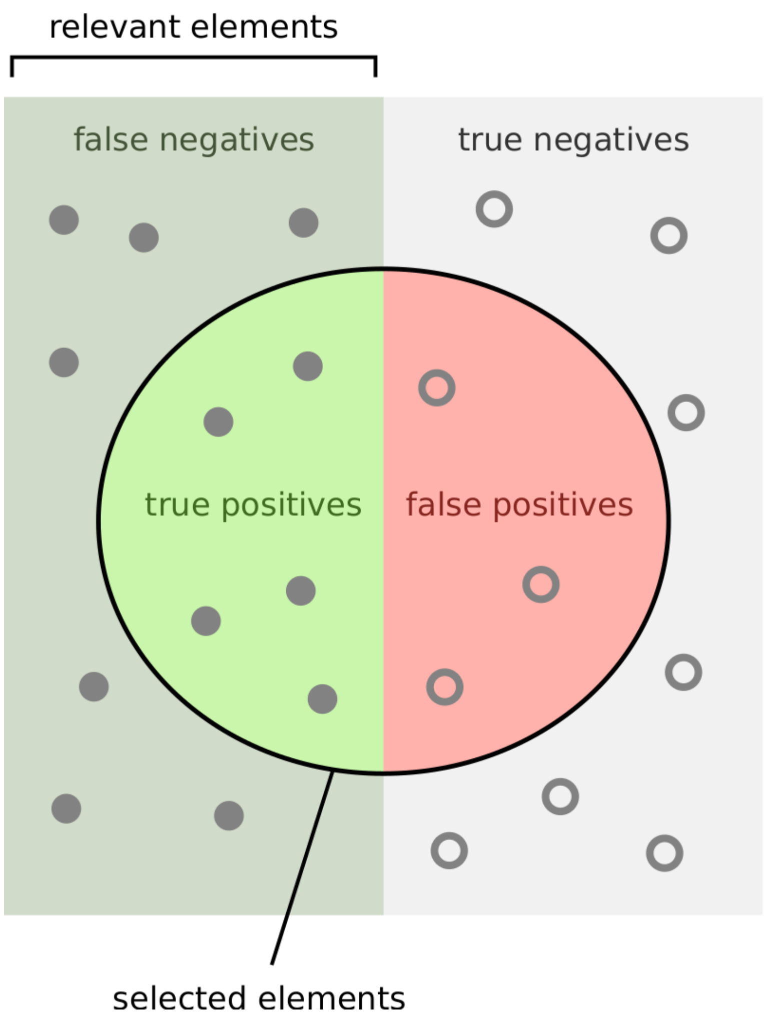
1. **Naive Bayes:** Takes features as independent signals and fits the label according to Bayes Rule
2. **Support Vector Machine:** Finds a separator given a (non-)linear combination of features
3. **Random Forest:** Finds hierarchical decision rules that divide the texts in classes

In supervised sentiment analysis, generating the ground truth data is the most critical part and is required to train the model. Producing sufficient annotations from readers or authors can be expensive.

Supervised methods are usually not out-of-the-box like unsupervised tools, you would have to fit your own model to a ground truth dataset.

Evaluating classifiers

- True positives TP: All positive cases that are correctly predicted
- False positives FP: All negatives that were wrongly predicted as positive
- True negatives TN: All negative cases that are correctly predicted
- False negatives FN: All positive cases that were incorrectly predicted as negative



Accuracy, Precision, and Recall

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

The measure of precision answers the question “How sure am I of this prediction?”

The measure of recall answers the question “How many of the things that I’m looking for are found?”

Balancing precision and recall

A way to compute a trade-off between Precision and Recall is the F_1 score, which is a harmonic mean of Precision and Recall:

$$F_1 = 2 * \frac{Precision * Recall}{Precision + Recall}$$

The F_1 score is often used as a way to summarize the quality of classifiers. When more than one class is possible, you should look at the mean of F_1 over the classess or to the F_1 of each class separately. The F_1 score is often used in sentiment analysis competitions to chose the best tools, for example in the SemEval 2017 competition.

Let someone else do it: Black-box APIs

Watson — Natural Language Understanding

Sentiment | Emotion | Keywords | Entities | Categories | Concept | Semantic Roles

Review the overall sentiment and targeted sentiment of the content.

JSON


```
{
 "sentiment": {
  "document": {
   "score": -0.74758,
   "label": "negative"
  }
 }
}
```


Overall Sentiment

Negative -0.75

Google Cloud Platform

A sample `analyzeSentiment` response to the Gettysburg Address is shown below:


```
{
  "documentSentiment": {
    "score": 0.2,
    "magnitude": 3.6
  },
  "language": "en",
   "sentences": [
    {
      "text": {
        "content": "Four score and seven years ago our fathers brought forth
        on this continent a new nation, conceived in liberty and dedicated to
        the proposition that all men are created equal.",
        "beginOffset": 0
      },
      "sentiment": {
        "magnitude": 0.8,
        "score": 0.8
      }
    },
```


Easy to use but data and methods unknown. Do your own evaluation!