

Turing School of Software and Design

Module 4 Capstone Project

August 2017

*Wakaru*

Customer Relationship Management Tone Analyzer

[Carl Richmond](https://www.linkedin.com/in/carl-richmond/)

[Repository](https://github.com/ACC25/wakaru)

Abstract

Curating user experiences has become central to many industries in the digital age, where presenting a unified story, across company and product, is central to a holistic marketing strategy (Zeiser 316). Customer Service represents one of the pillars of business and controlling a user’s experience in that area is important to presenting a unified front.

With the emergence of accessible natural language processing technologies, such as IBM’s Watson Tone Analyzer and Google’s Natural Language API, companies can now track the language and tone of their front-facing employees — such as customer service agents.

The present work is focused on developing methodologies to manage the data received from these natural language processing programs, in such a way as to infer whether front-facing employees are contributing to the story of a business or hurting its overall brand.

Wakaru, the tone analyzer app, aims to present these conclusions to businesses in an easy to understand manner.

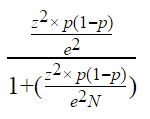
# Introduction

The history of natural language processing began in the early 1950s with Alan Turing proposing, in his article *Computing Machinery and Intelligence*, that one method to determine a machines intelligence would be to play an imitation game, where the computer must fool an interrogator into believing it is human. At a fundamental level this involves a program breaking down human sentences and text into machine understandable data.

The work around Wakaru does not involve contributing any new ideas to the field of natural language processing, but rather how a business can deal with the resulting data in a statistically sound way. It aims to answer the question of whether a business can use the data from IBM Watson’s API to inform its practices in Customer Service departments.

# Sample Selection

The sample size for Wakaru was determined using the following equation for a company of 200 employees, confidence level of 80% and margin of error of 5%.



Population size = N | Margin of error = e | Z-score = z

To reach the goal of 91 samples, surveys were sent out and recipients were prompted to reply as if they were customer service agents. All the surveys and results can be found [here](https://github.com/ACC25/wakaru). The surveys were split into several categories to focus the language of the respondents and help identify trends in different customer service interactions. Twelve different surveys were created, each with a unique prompt, to cover a variety of different scenarios.

Following the study done by Software Advice (Software Advice), the categories were set as: granting requests with a good tone, denying requests with a good tone and denying request with a bad tone. The study found that customer’s perception of tone changed dramatically between granting and denying interactions. Specifically, customers prefer a casual tone when being granted a request, but find it offensive when being denied one. The categories of the surveys, as with the prompts, were designed to focus in on the tone differences between granting and denying interactions.

# IBM Watson Tone Analyzer

IBM Watson’s Tone Analyzer returns 10 metrics per segment of text that you provide: disgust, fear, joy, sadness, anger, openness, conscientiousness, extraversion, agreeableness and emotional range. Each metric has a score between 0 (lowest) and x (highest). By themselves, they were found to mean very little – especially with professional language. The words we use in casual conversations are much different than the words we use in a professional conversation, even a casual one, and the tone scores indicate difficulty in classifying professional interactions.

For instance, many good toned responses taken from the sample surveys scored as low as bad toned responses across IBM’s metrics. Disgust for instance, had a STDEV of 0.02 across all categories, which indicates the metric is statistically insignificant and cannot help inform whether an interaction was ultimately good or bad – by itself. Anger, fear, openness and emotional range had similar results. The lack of deviation across the tonal categories highlights the difficulty in classifying language and the need to look at all the results together, and focus on how the scores correlate to each other.

# Classification

To differentiate the good emails from the bad, the tonal categories with the strongest correlations were combined into scores. The Pearson Correlation Coefficient or Bivariate Correlation were used to rank the relationships between the different tonal categories.



Through linear regressions and looking at the STDEV of all the categories, Joy and Sadness stood out as being strongly correlated to the categories of the emails, while also being moderately negatively correlated to each other. Their relationship with one another, calculated using the Pearson Correlation, and their other strongest relationships are displayed below.

../../../../Desktop/breakdown.jpeg

# Bibliography

Software Advice. "The Best Tone for Email Customer Support." 2014. *Software Advice.* <http://www.softwareadvice.com/resources/the-best-tone-for-email-customer-support/>.

Zeiser, Anne. *Transmedia Marketing: From Film and TV to Games and Digital Media*. New York: Focal Press, 2015.