

Brain Language Metrics on Earnings Call Transcripts

Product Summary

The exploitation of textual unstructured content (news, company filings, earnings calls etc) in financial analysis is quickly expanding across both quantitative and discretionary strategies as demonstrated by the growing number of academic papers and products in this domain.

The Brain Language Metrics on Earnings Calls Transcripts (BLMECT) dataset has the objective of monitoring several language metrics for the quarterly earnings call transcripts of 4500+ US stocks.

The dataset is made of two parts; one includes the **language metrics** for the most recent earnings call transcript for each stock, namely:

1. Financial sentiment
2. Percentage of words belonging to financial domain classified by language types:
 - “Constraining” language
 - “Litigious” language
 - “Uncertainty” language
3. Readability score
4. Lexical metrics such as lexical density and richness
5. Text statistics such as the transcript length

The second part includes the **differences** between the most recent earnings call transcript and the previous one:

1. Difference of the various language metrics (e.g. delta sentiment, delta readability score delta, delta percentage of a specific language type etc.)
2. Similarity metrics between documents, also with respect to a specific language type (for example similarity with respect to “litigious” language or “uncertainty” language)

The metrics calculation is reported separately for the following **sections** of the transcript:

- a. Management Discussion
- b. Analysts Questions
- c. Management Answers to Analysts Questions

Dataset Frequency

The dataset is updated with a daily frequency since new earnings calls transcripts are published every day for some of the universe stocks. Clearly the data for each stock will change on a quarterly basis when new earnings calls are published. The historical dataset is available from year 2012.

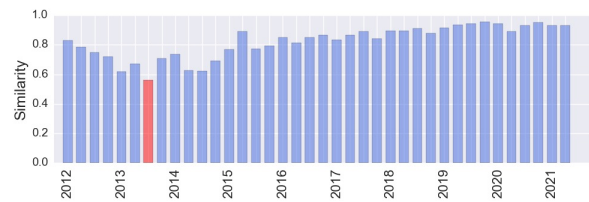
Last update: 07/07/2021

DISCLAIMER: the content of this presentation is not to be intended as investment advice. The material is provided for informational purposes only and does not constitute an offer to sell, a solicitation to buy, or a recommendation or endorsement for any security or strategy, nor does it constitute an offer to provide investment advisory or other services by Brain. Brain makes no guarantees regarding the accuracy and completeness of the information expressed in this document.

Examples of Metrics

One of the dataset metrics is the similarity between the most recent earnings call transcript and the previous one for various sections of the transcript.

Similarity with focus on generic financial domain language (“Management Discussion” section) - MSFT



Similarity with focus on “uncertainty” financial domain language (“Management Discussion” section) - MSFT



Another metric included in the dataset is the difference of sentiment score measuring the tone of the Analyst Questions as shown below.

Difference of Sentiment Score in “Analyst Questions” in AAPL Earnings Calls Transcripts.



Contacts

BRAIN is a Research Company that develops proprietary signals based on alternative data and algorithms for investment strategies on financial markets.

- EMAIL: contact@braincompany.co
- WEB: <http://www.braincompany.co>