BRAIN

# Brain Language Metrics on Earnings Calls Transcripts – V2.0

*Last Update 24/07/2025*

## File "stock_list_$date.csv"

The file contains the full list of stocks for which the earnings call transcripts are monitored daily for a specific date. The file is marked with the export date *$date.*

| Field | Type | Description | Example |
|---|---|---|---|
| COMPOSITE_FIGI | String | The FIGI composite code (https://www.openfigi.com) that uniquely identifies the stock across related exchanges in the same country | BBG000CL9VN6 |
| TICKER | String | The stock ticker | NFLX |
| NAME | String | The company name | Netflix Inc. |
| SECTOR | String | The company sector | communication services |

## File "metrics_earnings_calls_*$date*.csv"

The file contains the **language metrics** calculated on the earnings call transcripts for a specific date. The file is marked with the export date *$date* and it is released with daily frequency within 10 AM UTC.

| Field | Type | Range | Description | Example |
|---|---|---|---|---|
| COMPOSITE_FIGI | String | - | The FIGI composite code (https://www.openfigi.com) that uniquely identifies the stock across related exchanges in the same country | BBG000CL9VN6 |
| TICKER | String | - | The stock ticker | NFLX |
| DATE | String | - | The calculation date in YYYY-MM-DD format. The updated file for the current DATE is published every day within 10AM UTC. | 2021-02-01 |
| LAST_TRANSCRIPT_DATE | String | - | The date of last earnings call transcript (with respect to DATE) issued by the company in YYYY-MM-DD format. | 2021-01-20 |
| LAST_TRANSCRIPT_QUARTER | Number | [1,4] | Reference quarter of last earnings call transcript | 4 |
| LAST_TRANSCRIPT_YEAR | Number | [0,inf] | Reference year of last earnings call transcript | 2020 |
| In the following, we present a set of language metrics calculated for three sections of the earnings call transcript:<br>• **Management Discussion (MD)**: Contains statements from management presenting the company's results during the initial part of the call.<br>• **Analysts' Questions (AQ)**: Includes questions asked by analysts during the Q&A portion of the call. | | | | |

- **Management Answers (MA)**: Includes responses provided by executives to analysts' questions during the Q&A session.

**Notes:**

- If the fields are empty, it means that the extraction of the specific section from the transcript failed (e.g., the section was missing), or the extracted text was insufficient to compute meaningful metrics.
- The number of calculated metrics may vary by section. For example, 9 metrics are computed for the MD section, while only 5 may be available for the AQ section.

| | | | | |
|---|---|---|---|---|
| MD_N_CHARACTERS | Number | [0, Inf] | The length of the section "Management Discussion" measured in number of characters. | 17300 |
| MD_SENTIMENT | Number | [-1,+1] | The financial sentiment for the "Management Discussion" section of the last available transcript. | 0.25 |
| MD_SCORE_UNCERTAINTY | Number | [0, 1] | The percentage of financial domain "*uncertainty"* language present in the "Management Discussion" section of the last transcript. | 0.15 |
| MD_SCORE_LITIGIOUS | Number | [0, 1] | The percentage of financial domain *"litigious"* language present in the "Management Discussion" section of the last transcript. | 0.15 |
| MD_SCORE_CONSTRAINING | Number | [0, 1] | The percentage of financial domain *"constraining"* language present in the "Management Discussion" section of the last transcript. | 0.1 |
| MD_READABILITY | Number | [0, Inf] | Reading grade level for the "Management Discussion" section of the last available transcript. | 18.2 |
| MD_LEXICAL_RICHNESS | Number | [0, 1] | Lexical richness for the "Management Discussion" section of the last available transcript. | 0.2 |
| MD_LEXICAL_DENSITY | Number | [0, 1] | Lexical density for the "Management Discussion" section of the last available transcript. | 0.5 |
| MD_SPECIFIC_DENSITY | Number | [0, 1] | Percentage of words belonging to the specific dictionary used for the earnings call analysis present in the "Management Discussion" section of the last available transcript. | 0.1 |
| AQ_N_CHARACTERS | Number | [0, Inf] | The length of the section "Analyst Questions" measured in number of characters. | 17300 |
| AQ_SENTIMENT | Number | [-1,+1] | The financial sentiment for the section "Analyst Questions" of the last available transcript. | 0.25 |
| AQ_SCORE_UNCERTAINTY | Number | [0, 1] | The percentage of financial domain "*uncertainty"* language present in the section "Analyst Questions" of the last transcript. | 0.15 |

| | | | | |
|---|---|---|---|---|
| AQ_SCORE_LITIGIOUS | Number | [0, 1] | The percentage of financial domain *"litigious"* language present in the section "Analyst Questions" of the last transcript. | 0.15 |
| AQ_SCORE_CONSTRAINING | Number | [0, 1] | The percentage of financial domain *"constraining"* language present in the section "Analyst Questions" of the last transcript. | 0.1 |
| MA_N_CHARACTERS | Number | [0, Inf] | The length of the section "Management Answers" measured in number of characters. | 17300 |
| MA_SENTIMENT | Number | [-1,+1] | The financial sentiment for the "Management Answers" section of the last available transcript. | 0.25 |
| MA_SCORE_UNCERTAINTY | Number | [0, 1] | The percentage of financial domain "*uncertainty"* language present in the "Management Answers" section of the last transcript. | 0.15 |
| MA_SCORE_LITIGIOUS | Number | [0, 1] | The percentage of financial domain *"litigious"* language present in the "Management Answers" section of the last transcript. | 0.15 |
| MA_SCORE_CONSTRAINING | Number | [0, 1] | The percentage of financial domain *"constraining"* language present in the "Management Answers" section of the last transcript. | 0.1 |
| MA_READABILITY | Number | [0, Inf] | Reading grade level for the "Management Answers" section of the last available transcript. | 18.2 |
| MA_LEXICAL_RICHNESS | Number | [0, 1] | Lexical richness for the "Management Answers" section of the last available transcript. | 0.2 |
| MA_LEXICAL_DENSITY | Number | [0, 1] | Lexical density for the "Management Answers" section of the last available transcript. | 0.5 |
| MA_SPECIFIC_DENSITY | Number | [0, 1] | Percentage of words belonging to the specific dictionary used for company filings analysis present in the "Management Answers" section of the last available transcript. | 0.1 |

## File "differences_earnings_call_*$date*.csv"

The file contains the **similarity measures and differences of the language metrics** between the two last earnings call transcripts (last quarter Q and previous quarter Q-1) available for each stock on the specific export *$date.*
For example for *$date* = 1st February 2021 for stock NFLX the file will contain the similarity measures and differences of metrics between the

- Last available earnings call "Q4 – 2020" that has been published on 20th January 2021
- Previous earnings call transcripts "Q3 – 2020" that has been published on 21st October 2020

The file is marked with the export date *$date* and it is released with daily frequency within 10 AM UTC.

| Field | Type | Range | Description | Example |
|---|---|---|---|---|
| COMPOSITE_FIGI | String | - | The FIGI composite code (https://www.openfigi.com) that uniquely identifies the stock across related exchanges in the same country. | BBG000CL9VN6 |
| TICKER | String | - | The stock ticker | NFLX |
| DATE | String | - | The calculation date in YYYY-MM-DD format. The updated file for the current DATE is published every day within 10AM UTC. | 2021-02-01 |
| LAST_TRANSCRIPT_DATE | String | - | The date of last earnings call transcript (with respect to DATE) issued by the company in YYYY-MM-DD format. | 2021-01-28 |
| LAST_TRANSCRIPT_QUARTER | Number | [1,4] | Reference quarter of last earnings call transcript | 4 |
| LAST_TRANSCRIPT_YEAR | Number | [0,inf] | Reference year of last earnings call transcript | 2020 |
| PREV_TRANSCRIPT_DATE | String | - | The date of previous transcript (with respect to LAST_TRANSCRIPT_DATE) issued by the company in YYYY-MM-DD format. | 2020-10-21 |
| PREV_TRANSCRIPT_QUARTER | Number | [1,4] | Reference year of previous earnings call transcript (with respect to LAST_TRANSCRIPT_DATE) issued by the company. | 3 |
| PREV_TRANSCRIPT_YEAR | Number | [0,inf] | Reference year of previous earnings call transcript (with respect to LAST_TRANSCRIPT_DATE) issued by the company. | 2020 |

In the following, we present a series of similarity measures and differences in language metrics calculated between the most recent earnings call transcript and the previous one, for the following sections:

- **Management Discussion (MD)**: Contains executive statements presenting the company's results in the initial part of the call.
- **Analysts' Questions (AQ)**: Includes questions posed by analysts during the Q&A session.
- **Management Answers (MA)**: Includes responses provided by executives to analysts' questions in the Q&A session.

**Notes:**

| | | | | |
|---|---|---|---|---|
| • If the fields are empty, this indicates that the extraction of the specific section from the transcript failed (e.g., the section was missing) or that the text was insufficient to compute meaningful metrics.<br>• The number of calculated fields may vary by section. For example, the MD section includes 16 fields, while the AQ section includes only 5. | | | | |
| MD_DELTA_PERC_N_CHARACTERS | Number | [-Inf,+Inf] | Percentage change of length (measured in number of characters) between the "Management Discussion" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.1 |
| MD_DELTA_SENTIMENT | Number | [-2,+2] | The difference of financial sentiment between the "Management Discussion" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | -0.1 |
| MD_DELTA_SCORE_UNCERTAINTY | Number | [-1,+1] | The difference of percentage of financial domain "*uncertainty*" language between the "Management Discussion" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | -0.1 |
| MD_DELTA_SCORE_LITIGIOUS | Number | [-1,+1] | The difference of percentage of financial domain "*litigious*" language between the "Management Discussion" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | -0.1 |
| MD_DELTA_SCORE_CONSTRAINING | Number | [-1,+1] | The difference of percentage of financial domain "*constraining*" language between the "Management Discussion" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | -0.1 |
| MD_DELTA_READABILITY | Number | [-Inf, Inf] | The difference of the readability metric between the "Management Discussion" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | -0.5 |
| MD_DELTA_LEXICAL_RICHNESS | Number | [-1,+1] | The difference of the lexical richness metric between the "Management | 0.05 |

| | | | | |
|---|---|---|---|---|
| | | | Discussion" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | |
| MD_DELTA_LEXICAL_DENSITY | Number | [-1,+1] | The difference of the lexical density metric between the "Management Discussion" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.05 |
| MD_DELTA_SPECIFIC_DENSITY | Number | [-1,+1] | The difference of the specific density metric between the "Management Discussion" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.05 |
| MD_SIMILARITY_ALL | Number | [0, 1] | The language similarity between the "Management Discussion" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.8 |
| MD_SIMILARITY_POSITIVE | Number | [0, 1] | The similarity in terms of financial domain "*positive*" language between the "Management Discussion" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.8 |
| MD_SIMILARITY_NEGATIVE | Number | [0, 1] | The similarity in terms of financial domain "*negative*" language between the "Management Discussion" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.8 |
| MD_SIMILARITY_UNCERTAINTY | Number | [0, 1] | The similarity in terms of financial domain "*uncertainty*" language between the "Management Discussion" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.8 |
| MD_SIMILARITY_LITIGIOUS | Number | [0, 1] | The similarity in terms of financial domain "*litigious*" language between the "Management Discussion" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the | 0.8 |

| | | | | |
|---|---|---|---|---|
| | | | same section of the previous transcript (PREV_TRANSCRIPT_DATE). | |
| MD_SIMILARITY_CONSTRAINING | Number | [0, 1] | The similarity in terms of financial domain "*constraining*" language between the "Management Discussion" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.8 |
| AQ_DELTA_PERC_N_CHARACTERS | Number | [-Inf,+Inf] | Percentage change of length (measured in number of characters) between the "Analyst Questions" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.1 |
| AQ_DELTA_SENTIMENT | Number | [-2,+2] | The difference of financial sentiment between the "Analyst Questions" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | -0.1 |
| AQ_DELTA_SCORE_UNCERTAINTY | Number | [-1,+1] | The difference of percentage of financial domain "*uncertainty*" language between the "Analyst Questions" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | -0.1 |
| AQ_DELTA_SCORE_LITIGIOUS | Number | [-1,+1] | The difference of percentage of financial domain "*litigious*" language between the "Analyst Questions" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | -0.1 |
| AQ_DELTA_SCORE_CONSTRAINING | Number | [-1,+1] | The difference of percentage of financial domain "*constraining*" language between the "Analyst Questions" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | -0.1 |
| AQ_SIMILARITY_ALL | Number | [0, 1] | The language similarity between the "Analyst Questions" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous | 0.8 |

| | | | | |
|---|---|---|---|---|
| | | | transcript (PREV_TRANSCRIPT_DATE). | |
| AQ_SIMILARITY_POSITIVE | Number | [0, 1] | The similarity in terms of financial domain "*positive*" language between the "Analyst Questions" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.8 |
| AQ_SIMILARITY_NEGATIVE | Number | [0, 1] | The similarity in terms of financial domain "*negative*" language between the "Analyst Questions" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.8 |
| MA_DELTA_PERC_N_CHARACTERS | Number | [-Inf,+Inf] | Percentage change of length (measured in number of characters) between the "Management Answers" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.1 |
| MA_DELTA_SENTIMENT | Number | [-2,+2] | The difference of financial sentiment between the "Management Answers" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | -0.1 |
| MA_DELTA_SCORE_UNCERTAINTY | Number | [-1,+1] | The difference of percentage of financial domain "*uncertainty*" language between the "Management Answers" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | -0.1 |
| MA_DELTA_SCORE_LITIGIOUS | Number | [-1,+1] | The difference of percentage of financial domain "*litigious*" language between the "Management Answers" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | -0.1 |
| MA_DELTA_SCORE_CONSTRAINING | Number | [-1,+1] | The difference of percentage of financial domain "*constraining*" language between the "Management Answers" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | -0.1 |
| MA_DELTA_READABILITY | Number | [-Inf, Inf] | The difference of the readability metric between the "Management | -0.5 |

| | | | | |
|---|---|---|---|---|
| | | | Answers" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | |
| MA_DELTA_LEXICAL_RICHNESS | Number | [-1,+1] | The difference of the lexical richness metric between the "Management Answers" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.05 |
| MA_DELTA_LEXICAL_DENSITY | Number | [-1,+1] | The difference of the lexical density metric between the "Management Answers" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.05 |
| MA_DELTA_SPECIFIC_DENSITY | Number | [-1,+1] | The difference of the specific density metric between the "Management Answers" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.05 |
| MA_SIMILARITY_ALL | Number | [0, 1] | The language similarity between the "Management Answers" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.8 |
| MA_SIMILARITY_POSITIVE | Number | [0, 1] | The similarity in terms of financial domain "*positive*" language between the "Management Answers" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.8 |
| MA_SIMILARITY_NEGATIVE | Number | [0, 1] | The similarity in terms of financial domain "*negative*" language between the "Management Answers" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.8 |
| MA_SIMILARITY_UNCERTAINTY | Number | [0, 1] | The similarity in terms of financial domain "*uncertainty*" language between the "Management Answers" section of the last available transcript (LAST_TRANSCRIPT _DATE) and of the previous transcript of same period and category (PREV_TRANSCRIPT_DATE). | 0.8 |
| MA_SIMILARITY_LITIGIOUS | Number | [0, 1] | The similarity in terms of financial domain "*litigious*" language between the "Management Answers" section of the last available | 0.8 |

| | | | | |
|---|---|---|---|---|
| | | | transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | |
| MA_SIMILARITY_CONSTRAINING | Number | [0, 1] | The similarity in terms of financial domain "*constraining*" language between the "Management Answers" section of the last available transcript (LAST_TRANSCRIPT_DATE) and the same section of the previous transcript (PREV_TRANSCRIPT_DATE). | 0.8 |

## Contacts

For more information please contact *support@braincompany.co*

## Disclaimer