

# Gauge corporate performance by analysing earnings calls using NLP

---

Sreeram Ramesh

Public-company executives discuss corporate performance and expectations with analysts on earnings calls. With the help of technology, these calls are converted into transcripts for future reference. Earnings call transcripts provide vital information on a company's performance and market dynamics, and a glimpse of company executives' views. However, the unstructured nature of these transcripts has until recently made it difficult to analyse them in a scalable manner. The recent progress in financial data science, though, has made such analysis possible.

It is taxing for an analyst or portfolio manager to physically listen to every major conference call during earnings seasons. Going through all transcript reports manually is also not an efficient use of front-office bandwidth. This is where the power of natural language processing (NLP) comes into play. A completely automated NLP algorithm can extract text from transcript reports and perform the required analysis in a matter of seconds. This would rapidly increase the number of earnings call transcripts analysed and open the way for insights to be extracted.

## **An example of analysing text using NLP**

A text analytics solution built on NLP technology does much of the heavy lifting when it comes to gleaning valuable information from earnings calls. The current NLP solution pipeline is as follows:

- Extract text data from PDF files covering earnings calls
  - Identify management participants and covering analysts
  - Separate comments made by management and analysts
  - Extract specific themes and keywords
    - Gender-biased content
    - Industry-specific content
-

- Company-specific content
- Detect the gender of members of the management team
- Calculate the number of females and the number of males among company participants

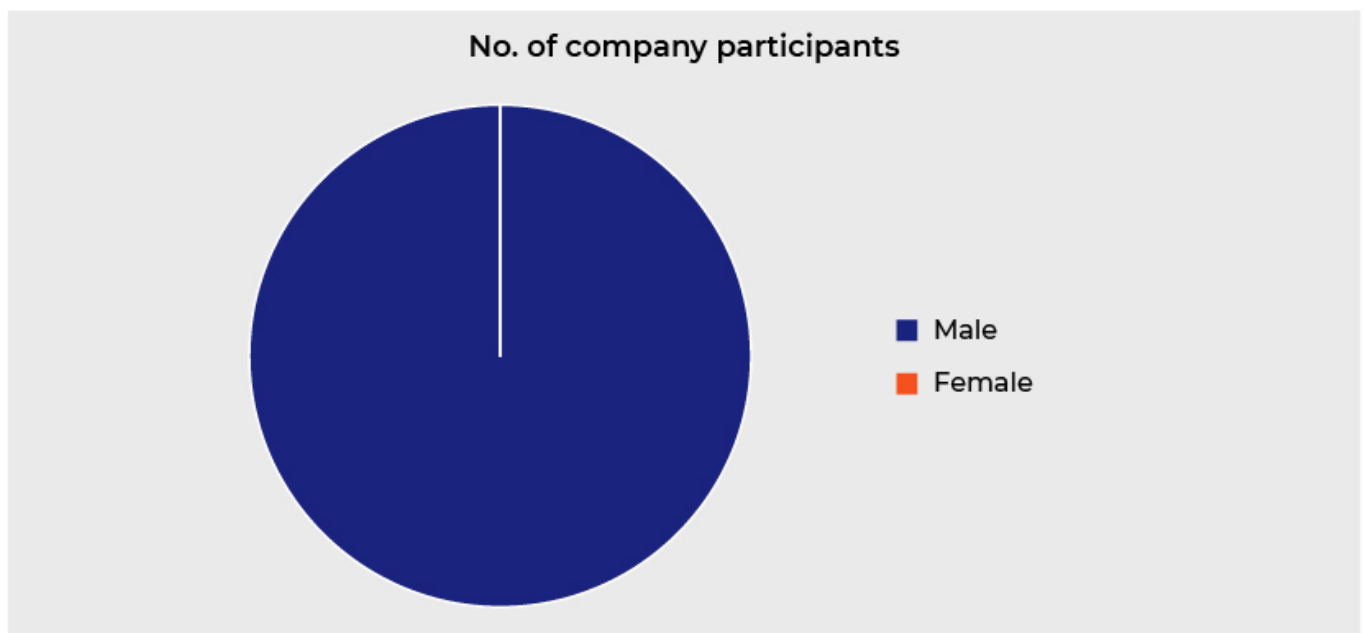
### **A deep dive into the analysis**

We use a combination of PDF-parsing libraries to extract textual information from the reports. Once the textual data is extracted, a completely automated NLP algorithm will extract company participants' and analysts' names from the report and separate the content by company participant and analyst.

We use state-of-the-art (SOTA) named entity-recognition models to extract the names from the company participants list and then use gender-detection algorithms to identify the gender of these individuals to calculate the number of females and the number of males.

Based on a predefined set of input keywords, the algorithm will automatically extract paragraphs containing the keywords. This will be done separately using the comments of the company participants and analysts.

### **Insight on the findings**

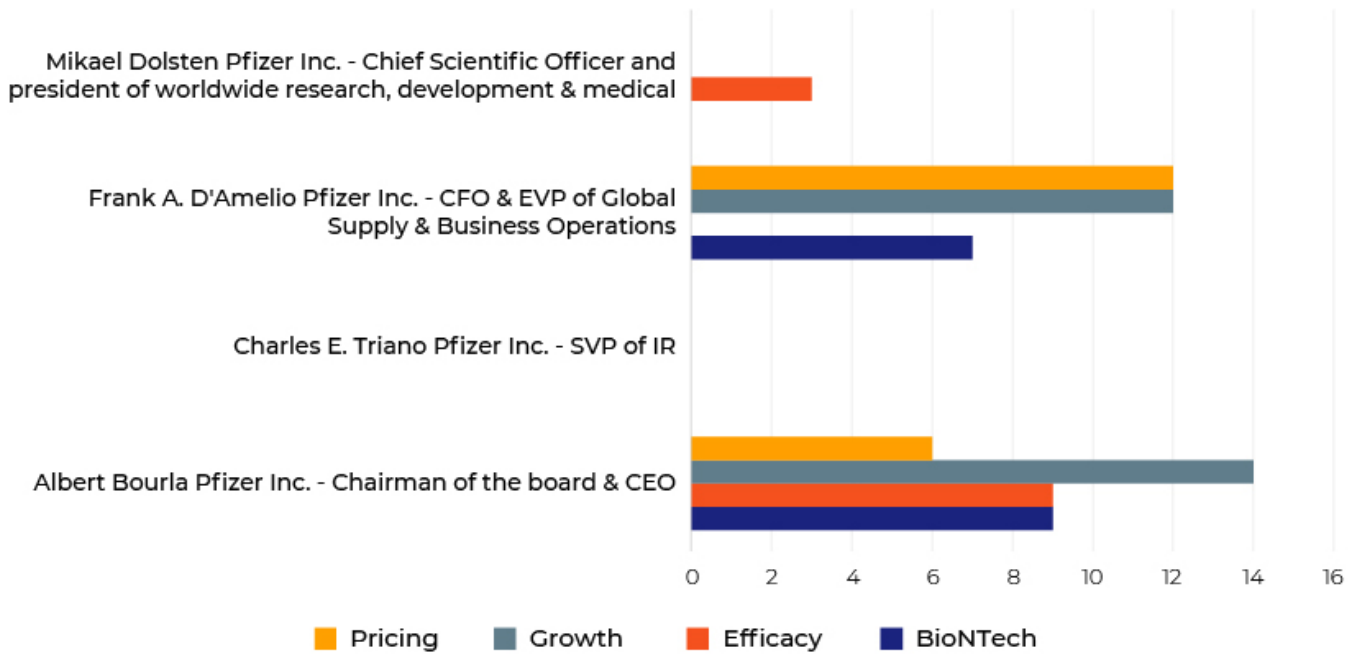


Source: Company reports

The number of female participants on the call helps gauge a company's diversity and inclusion policy. Pfizer's report shows only male company participants.

---

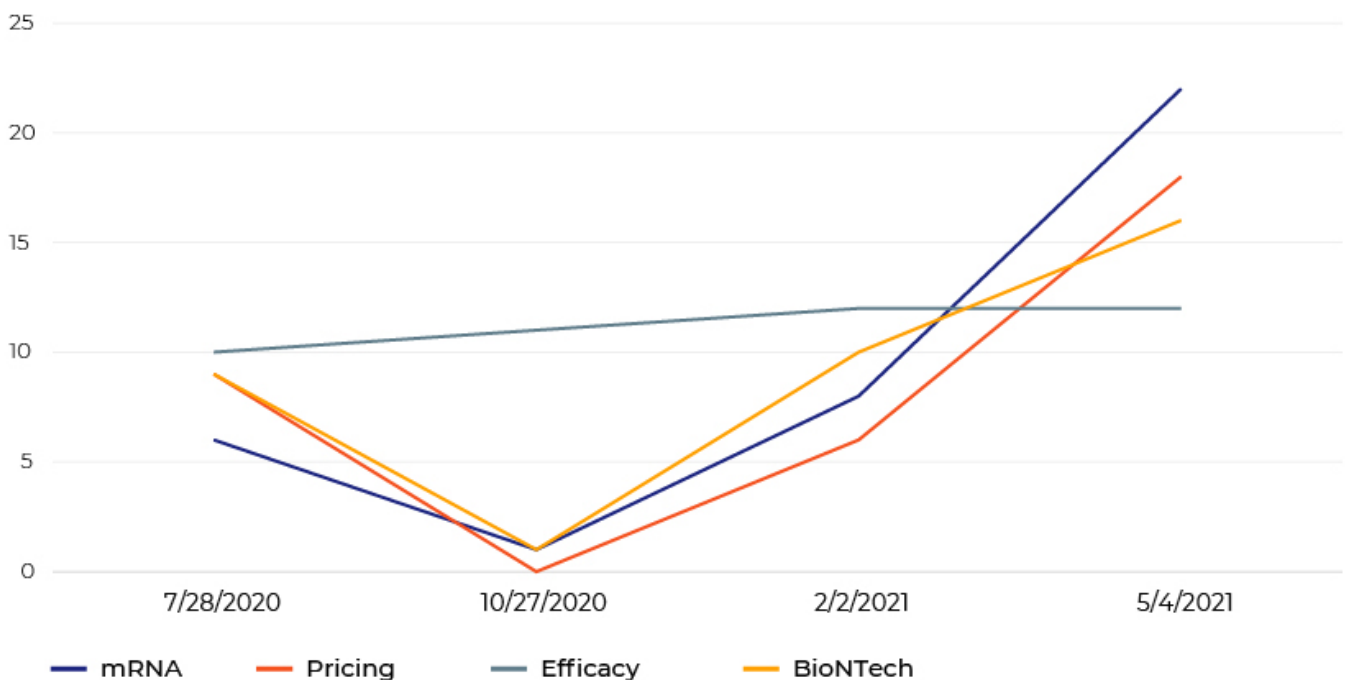
### Keywords used by company participants



Source: Company reports

The CEO used these keywords the most, and only the CEO and the CFO spoke of expected growth. The following graph illustrates findings from Pfizer’s transcript reports in recent quarters.

### Trend of keywords in Pfizer's transcript reports



Source: Company reports

The use of keywords BioNTech and mRNA has increased in the last two earnings calls; the use of “efficacy” has remained almost constant across calls. We have considered these words only for the purpose of illustration. The underlying code can consider any set of keywords, count how many times they are used and isolate the paragraphs containing the keywords.

Checking the use of gender-biased words such as “chairman”, “salesman” and “man-hours” could be

used to gauge a company's attitude towards diversity and inclusion.

### **Other areas that could be analysed using this method**

The extracted data could be used to generate more insight, which could be used to train multiple machine-learning (ML) models for a number of downstream tasks. The following are some interesting analyses that could be conducted:

- Sentiment analysis of answers given by management participants
- Report summary using pre-built summarisation algorithms
- Automatic extraction of new themes and events by comparing content over time

### **Conclusion**

With the advancements in technology and NLP techniques within financial data science, extracting deeper insights from large, unstructured datasets has become easier and quicker. Building ML and artificial intelligence (AI) models on top of the extracted data and insights makes the data far more powerful and useful. Similar analysis could be conducted on any company using text from internal publications, external communications and market news coverage.

Acuity Knowledge Partners can build bespoke NLP and AI algorithms to extract unstructured text from any source format (.pdf, .txt, images, etc.) and generate deeper insights based on the business context and problem. The solution can be completely automated along with a user-friendly and interactive front end as per the client requirement and data availability.

#### **About Acuity Knowledge Partners**

Acuity Knowledge Partners is a leading provider of high-value research, analytics and business intelligence to the financial services sector. The company supports over 400+ financial institutions and consulting companies through a team of over 4,000+ subject matter experts who work as an extension of the clients' teams based out of various global delivery centres.

We empower our clients to drive revenues higher. We innovate using our proprietary technology and automation solutions. We enable our clients to transform their operating model and cost base.