

Audio Sentiment Analysis

Term Project
Alexander Wang



1

Business Case



Scenario

- ▶ **Company:** Golden Lion Bank - Over 1000+ Investment Advisors
- ▶ **Sponsors:** Personal & Commercial Banking
- ▶ **Stakeholders:** Investment Advisors have to spend a lot of time doing research on companies to determine companies to invest in for their clients
- ▶ Including listening to earnings calls which can last up to 2 hours (for one company) - Market Sentiment
- ▶ **How can we help IAs save time by using NLP?**



“ Market sentiment refers to the **overall attitude of investors** toward **a particular security or financial market**. It is the **feeling or tone** of a market, or its **crowd psychology**, as revealed through the activity and price movement of the securities traded in that market. In broad terms, **rising prices** indicate **bullish market sentiment**, while **falling prices** indicate **bearish market sentiment**.



Finance

How WallStreetBets Pushed GameStop Shares to the Moon

By [Brandon Kochkodin](#)

January 25, 2021, 9:57 AM EST

- ▶ One trader turned \$53,566 into more than \$11 million
- ▶ The mania was years in the making and started as a value play

GameStop explodes



SOURCE: CNBC





Requirements

Scalable – Support for increased transaction volume, or increased compute power on demand

Cost effective – Less upfront cost (Low CAPEX)

Secure – Can set access permissions and support encryption

Low Maintenance – Reliable – Have failover and additional servers

Setup quickly (initial POC) – Without requiring too much time to configure and acquire equipment

Cloud vs On-Prem Matrix (Effort Required)

<u>Low to High Effort Required</u>	<u>Cloud</u>	<u>On-Prem</u>	<u>Recommendation</u>
Scalability	Low	High	Cloud
CAPEX	None	High	Cloud
OPEX	Medium	Medium	Tie
Security	Medium	High	Cloud
Maintenance	Low	High	Cloud
Setup	Low	High	Cloud

GCP vs AWS



Google Natural Language API (Sentiment)

- ▶ \$0.0001 up to 1M units (Monthly)
- ▶ 1 unit = 1000 characters
- ▶ No minimum (Free) with 0-5K units
- ▶ \$0.0010 per character

Google Speech to Text API (Standard)

- ▶ \$0.00040 (\$/second)
- ▶ Free within 0-60 minutes

Assumptions:

10,000 requests a month

Average recording length 60 minutes

AWS Comprehend (Sentiment)

- ▶ \$0.0001 up to 10M units (Monthly)
- ▶ 1 unit = 100 characters
- ▶ 3 unit = 300 characters min per request
- ▶ Free up to 50K units of text
- ▶ \$0.01 per character

AWS Transcribe

- ▶ \$0.00040 (\$/second)
- ▶ Free 60 minutes per month

GCP vs AWS



GCS Bucket

- ▶ 20.48 USD per Month (Standard) - Iowa US Central-1

Google App Engine

- ▶ F1 (1GB Memory)
- ▶ 0 USD (Includes Free)

AWS S3 Bucket

- ▶ 26.37 USD per Month (US East Ohio)

AWS Lambda

- ▶ 1GB Memory
- ▶ 0 USD (Includes Free)

Assumptions:

10,000 requests a month

Average recording length 60 minutes (100MB) - 1TB Total

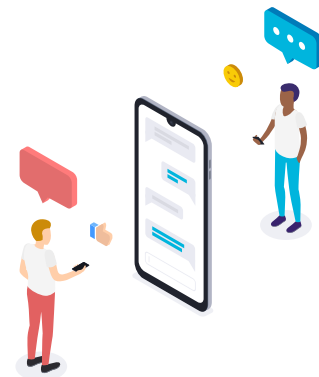
2

Solution Architecture



► Functionality

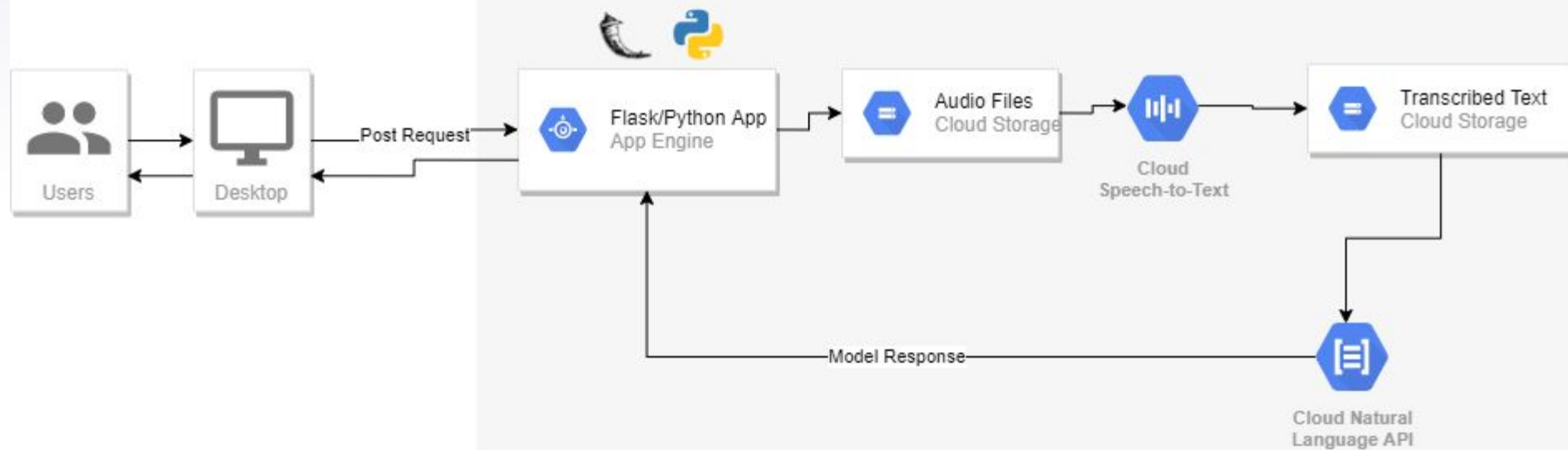
- ▶ **Need to transcribe audio recording**
- ▶ **Need to process the text and analyze the sentiment**
- ▶ **Web facing application which the user will upload an audio recording**
- ▶ **Hosting the web application**
- ▶ **Storing the raw text and store the audio files**



Technology

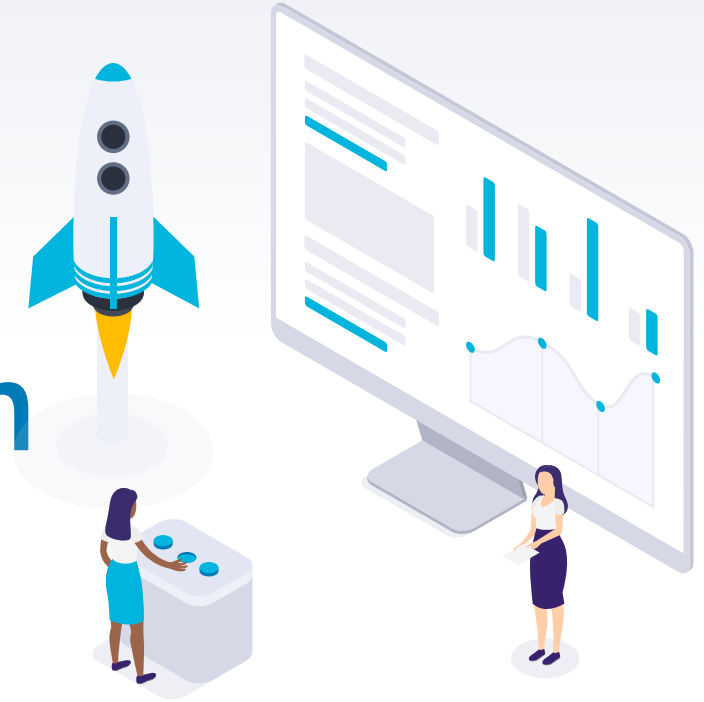
- ▶ **Speech to Text API - Audio Transcribe**
- ▶ **Natural Language API - Analyze Sentiment**
- ▶ **Flask/Python to serve the application to the user**
- ▶ **App Engine to host the application**
- ▶ **GCS Bucket (Standard) to store the text and audio files**





3

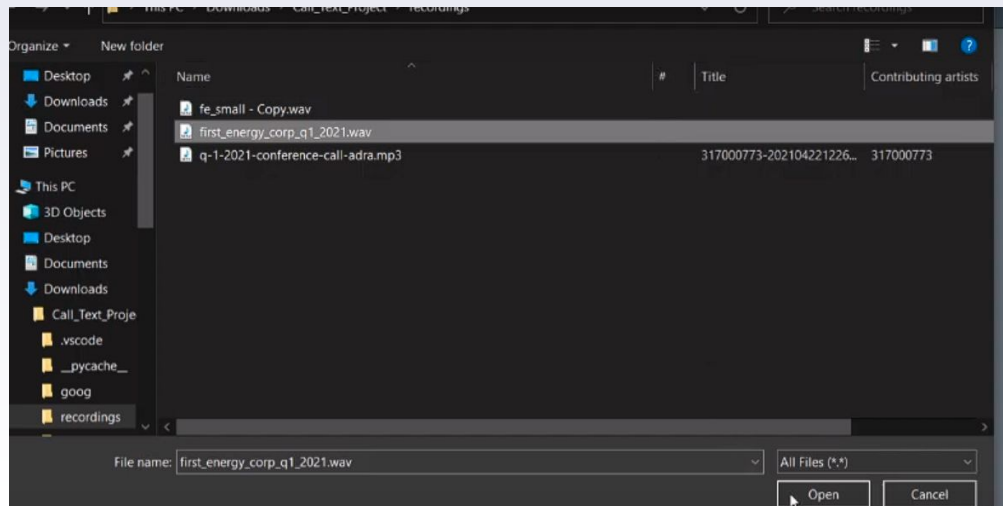
Demo - Solution



File Upload

Choose File No file chosen

Submit



File Upload

Choose File first_energy...1_2021.wav

Submit

```
"GET /" 200
first energy corp q1 2021.wav
Uploading...
Upload Complete
gs://audio records conference calls/first energy corp q1 2021.wav
Waiting for response to complete...
```

☰

Google Cloud Platform

call-text-analysis

Search products and resources

▼

🗨️

?

1

🔧 App Engine

Dashboard

Services

Versions

Instances

Versions

REFRESH

DELETE

STOP

START

MIGRATE TRAFFIC

SPLIT TRAFFIC

LEARN

SHOW INFO

Filter

Filter versions

<input type="checkbox"/>	Version	Status	Traffic Allocation	Instances ?	Runtime	Environment	Size ?	
<input type="checkbox"/>	20210426t010044 🔗	Serving	<div></div> 100%	0	python	Flexible	0 B	/v

☰

Google Cloud Platform

call-text-analysis

Search products and resources

▼

🗨️

🗄️ Cloud Storage

Browser

Monitoring

Settings

Browser

CREATE BUCKET

DELETE

REFRESH

Filter

Filter buckets

<input type="checkbox"/>	Name ↑	Created	Location type	Location	Default storage class ?
<input type="checkbox"/>	audio_records_conference_calls	Apr 24, 2021, 6:45:27 PM	Multi-region	us (multiple re...	Standard
<input type="checkbox"/>	audio_transcripts_raw	Apr 25, 2021, 8:32:27 PM	Multi-region	us (multiple re...	Standard

Google Cloud Platform

call-text-analysis

Search products and resources

Cloud Storage

Browser

Monitoring

Settings

Bucket details

REFRESH

LEARN

audio_records_conference_calls

OBJECTS

CONFIGURATION

PERMISSIONS

RETENTION

LIFECYCLE

Buckets > audio_records_conference_calls

UPLOAD FILES

UPLOAD FOLDER

CREATE FOLDER

MANAGE HOLDS

DOWNLOAD

DELETE

Filter by name prefix only

Filter

Filter objects and folders

<input type="checkbox"/>	Name	Size	Type	Created time ?	Storage class	Last modified	Public access ?	Encryption ?	
<input type="checkbox"/>	first_energy_corp_q1_2021.wav	26.7 MB	audio/x-wav	Apr 26, 2021, 7...	Standard	Apr 26, 20...	Not public	Google-managed key	

Google Cloud Platform

call-text-analysis

Search products and resources

Cloud Storage

Browser

Monitoring

Settings

Bucket details

REFRESH

LEARN

audio_transcripts_raw

OBJECTS

CONFIGURATION

PERMISSIONS

RETENTION

LIFECYCLE

Buckets > audio_transcripts_raw

UPLOAD FILES

UPLOAD FOLDER

CREATE FOLDER

MANAGE HOLDS

DOWNLOAD

DELETE

Filter by name prefix only

Filter

Filter objects and folders

<input type="checkbox"/>	Name	Size	Type	Created time ?	Storage class	Last modified	Public access ?	Encryption ?	
<input type="checkbox"/>	first_energy_corp_q1_2021.wav.txt	4.2 KB	text/plain	Apr 26, 2021, 7...	Standard	Apr 26, 20...	Not public	Google-managed key	

Sentiment Results

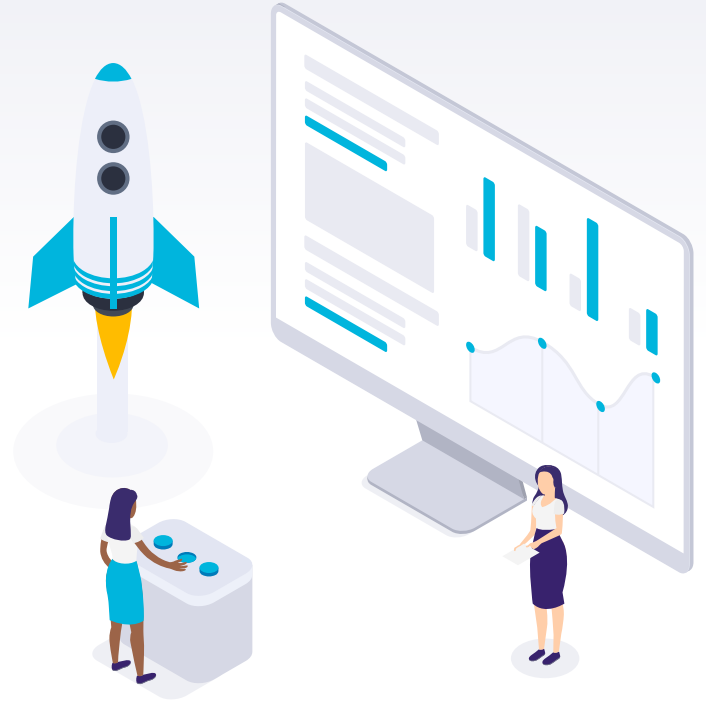
{ "Document sentiment score": 0.30000001192092896, "Document sentiment magnitude": 0.8999999761581421 }

Transcript of Recording

b"Springs and welcome to the First Energy Corp first quarter 2021 earnings conference call at this time all participants are in a listen-only mode a brief question-and-answer session will follow the formal presentation if anyone should require operator assistance during the conference please press star zero on your telephone keypad add reminder this conference is being recorded it's now my pleasure to introduce your host Irene Frizzell vice president investor relations for First Energy Corp thank you Miss pezzella May begin thank you good morning everyone and Welcome to our first quarter earnings call today we will make various forward-looking statements regarding revenues learning performance strategies Prospect and other matters these statements are based on current expectations and are subject to risks and uncertainties factors that could cause actual results to differ materially from those indicated by these statements can be found on the investor section of our website under the earnings information length and in our SEC filing we will also discuss certain non-gaap Financial measures reconciliation between gaap and non-gaap financial measures the presentation that supports today's discussion and other detailed information about the quarter can be found in the Strategic and financial highlights document on the investor section of our website participants in today's call include our president and Chief Executive Officer Steve straw senior vice president and Chief Financial Officer John Taylor and our vice chairperson and executive director John Somerhalder will also have several other Executives available to join us for the Q&A session now I'll turn the call over to Steve thank you Irene and good morning everyone yesterday we reported first-quarter 2021 gaap earnings of \$0.62 per share and operating earnings of \$0.69 per share which is in the upper end of our guidance range as John will discuss our results reflect the continued successful implementation of our investment strategies higher weather adjusted load in our residential class and strong financial discipline in managing our operating expenses last month I was honored to be named First Energy CEO and appointed to the board of directors I greatly appreciate the trust and confidence the board has placed in me since I was named president last met in the acting CEO in October I have great pride in First Energy and the work are employees due to serve our customers and communities it's my privilege to continue leading the company as we navigate our current challenges and position our business for long-term stability and success as we work to move First Energy forward my priorities are the continued safety of our employees and customers ensuring that ethics accountability and integrity are deeply ingrained in our culture and supported by a strong corporate compliance program executing Fe forward are transformational effort to capitalize on our potential deliver long-term results and maximize near-term Financial flexibility and continuing our investments in infrastructure growth opportunities from electrification grid modernization and renewable energy graichen to benefit our customers during today's call I'll provide an update on the Department of Justice investigation regulatory matters and RFP forward initiative and other business development John Somerhalder will join us for an update on the board and Management's work towards instilling a culture of compliance built upon the highest standards of ethics and integrity then John Taylor will review our results and other Financial topics before we open it up for your questions as we discussed on our fourth quarter call we are committed to taking decisive actions to rebuild a reputation and focus on the future and continuing to cooperate with the ongoing government investigations we have begun discussions with the doj regarding the resolution of this matter including the possibility First Energy entering into a deferred prosecution agreement we can't currently predict the timing outcome or the impact of the possible resolution with the doj our goal is to take a holistic and transparent approach with a range of stakeholders across the spectrum of matters under review this approach is consistent with the changes worm"

4

Final Thoughts

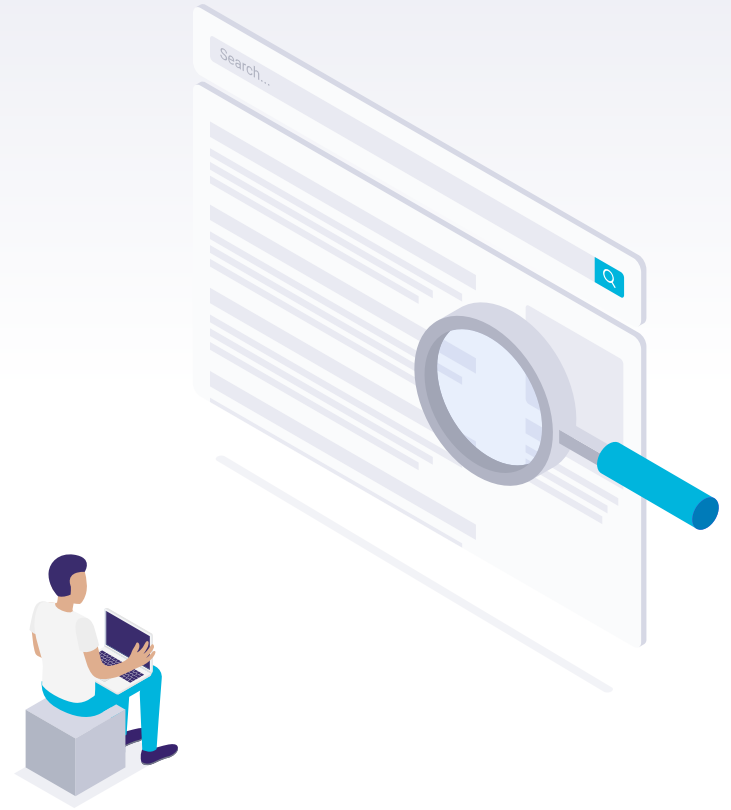


► Challenges & Lessons Learned

- ▶ Have to compromise between each providers - Look at from holistic perspective
- ▶ Overall solution rather than individual pieces
- ▶ Cost is not super intuitive to calculate (need to estimate demand)
- ▶ Local development != Cloud environment
- ▶ Many many different ways and tools to solve a problem

Thanks

Any questions?



► Appendices

YouTube (Demo)

<https://youtu.be/Nj25epGkF54>

Application Link

[https://call-text-analysis.uc.r.a
ppspot.com](https://call-text-analysis.uc.r.a
ppspot.com)

GitHub

[https://github.com/wangalexanders
k/call_text_project](https://github.com/wangalexanders
k/call_text_project)