



## PROJECT BACKGROUND



### THE PROBLEM

Scattered, jargon-filled information about the financial market makes it difficult to identify the right investment opportunity. There is a lack of integrated big data source & analytics.



### THE OBJECTIVE

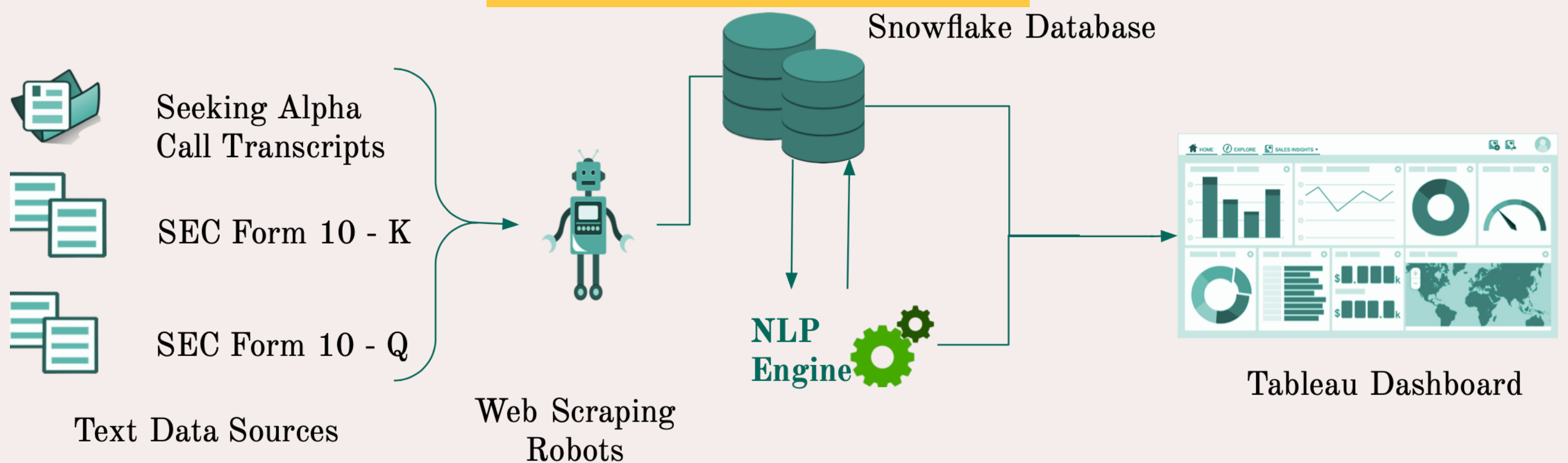
Design RPA data pipeline which integrates data from from 3 disparate sources over the year. This data source feeds into a NLP-powered Dashboard that provides time-series based sentiment & credibility analysis



### IMPACT & PRIORTITES

Provide competitive intelligence to portfolio managers, the internal team and customers. It will provide market interaction & trend analysis to drive investment intelligence using NLP

## PROJECT APPROACH



## ROBOTIC PROCESS AUTOMATION & NATURAL LANGUAGE PROCESSING

### RPA DATA PIPELINE

- Scalable UiPath Pipeline
- Integrated with Snowflake Database for storing information



Performed abstract summarization and topic modelings

Sentiment score and label for transcripts & filings to understand industry trends

Summarization

Sentiment analysis

Credibility analysis

These documents are intentionally complicated to hide information. Credibility score is generated by using four methods - gunning fog, flesch reading ease, dale, smog - to rate their credibility

## RESULTS & BUSINESS OUTCOME

### COMPREHENSIVE DATA SOURCE



Integrated 20+ years of filings and transcripts of 500 companies in clean, normalized raw & analytical databases

### AUTOMATED NLP ANALYSIS



Automated NLP analysis using python to extract insights from documents to analyze industry trends and market interactions

### TABLEAU DASHBOARD



Portfolio Managers, Investment Bankers & Analysts can interact with the Tableau dashboard to gain market intelligence