

Alex Abraham
146 McGlynn Place, Cedarhurst NY 11516
(201) 403-7591 • alexgabraham1@gmail.com
<https://alexabraham.dev>

Education

Georgia Tech, Online Masters in Computer Science

- Master of Science – Expected May 2023
- Major: Computer Science
 - Specialization: Computing Systems

Yeshiva University, Yeshiva College, New York, NY

- Bachelor of Arts – May 2016
- Major: Computer Science
 - Gertrude Nissenbaum Memorial Award for Excellence in Computer Science

Technical Skills

- Programming Languages (Python, Java, JavaScript, Go)
- Databases (MySQL, MongoDB, Snowflake, DBT)
- Cloud Infrastructure (GCP, Azure, AWS)
- Kafka, Apache Airflow, Kubernetes

Experience

Sr. Data Engineer – Data Ops, Vimeo, New York, NY

August 2022 – Present

- Designed and developed entity-based data model for product analysis
 - Entities were designed to improve data governance and discoverability
 - Created data standards by enforcing entity schemas cross-product

Data Engineer – Data Ops, Vimeo, New York, NY

October 2020 – August 2022

- Worked with product leaders across the company on event pipeline creation
- Used insights from working with product teams to improve our underlying data infrastructure

Data Engineer – Data Platform, Vimeo, New York, NY

June 2019 – October 2020

- Developed infrastructure for data ingestion pipelines across all Vimeo product offerings
- Built ETL pipelines for parsing CDN logs

Data Engineer, AllianceBernstein, New York, NY

July 2017 – June 2019

- Created robust data pipelines for analyzing alternative data
 - Collected data through web scraping (Scrapy + Splash, Selenium)
 - Performed analysis using Natural Language Processing
 - Integrated machine learning models developed by data scientists on the team
 - Utilized cloud technology to deploy pipelines
- Developed multiple web-based dashboards using modern web development techniques for visualizing model output (Angular front-end & flask back-end)
- Established team infrastructure in Microsoft Azure for large-scale data analysis
- Performed exploratory data analysis to initiate model development
- See [Appendix](#) for a summary on some of my projects

Software Developer, AllianceBernstein, New York, NY

July 2016 – July 2017

- Utilized web scraping techniques to build a new dataset for use in bond analysis
- Involved in redesigning process to calculate internal bond rating using data from multiple vendors
- Created processes to migrate data across internal systems

Appendix – Data Engineering Projects at AllianceBernstein

Client Churn

January 2019 – June 2019

AllianceBernstein has had very high client retention over the past decade. Our clients feel that their money is secure in our hands and trust us to make smart investment decisions on their behalf. That being said, there is still a need to best understand what drives certain clients to leave the firm.

- Performed Exploratory Data Analysis on internal data to help build a compound signal for client dropout
 - Incorporated dozens of unique data points on a monthly basis
- Started process to develop early warning indicator for advisors when their clients are at risk of leaving the firm

Earnings Transcripts

September 2018 – June 2019

Earnings Transcripts can prove to be a very useful alternative dataset when utilized properly. Every three months, company executives update the public on how their companies are performing. Comparing historical calls as well as calls with related companies can give insight as to how a company is doing comparatively.

- Used Selenium to scrape earnings transcripts from the internet
- Built data pipeline to retrieve latest earnings transcripts every day
- Developed a web-based dashboard to quickly search through transcripts for specific key words
 - Searches for a company across time as well as a list of companies over a specific quarter
- Integrated machine learning models for predicting company performance

SEC Edgar Database

July 2018 – December 2018

The SEC makes corporate filings publicly available on the EDGAR website. These filings contain valuable information which can be used to help predict the future success of a company. As the data engineer on the team, I created a data pipeline to process the filings from EDGAR.

- Used Scrapy (a Python based web scraping framework) to systematically retrieve documents from EDGAR
- Converted HTML documents from EDGAR to plaintext
- Performed additional data wrangling and formatting
- Incorporated machine learning model to analyze documents using Natural Language Processing
- Productionized pipeline for clear and consistent execution on a daily basis

Consumer Reviews

March 2018 – July 2018

Research teams at AllianceBernstein across multiple disciplines were curious about the possibility to develop a model using consumer reviews. When analyzed properly, consumer reviews can provide insight into the performance of products before official sales numbers are released.

- Built multiple web scrapers to retrieve consumer reviews from various websites
- Designed data schema to accommodate reviews from different websites
- Built front-end Angular dashboard allowing users to compare reviews across different car models and websites
 - Contains interactive charts with review statistics
 - Refreshed daily by tapping directly into data pipeline
- Used natural language processing to extract sentiment from review text

Short Sellers

November 2017 – January 2018

The goal of the short seller project was to determine the validity of short seller performance claims. Many short sellers announce their successes but it's hard to determine what percentage of their shorts are successful.

- Built web scrapers to extract data from social media and news outlets
- Created process to identify historical short targets
- Compared list of short targets with short seller claims