# **NFT Rating Project**

Naya DE Project by

Anton Vaysberg

Avital Chissick

Lev Rosenfeld

**Business idea:**

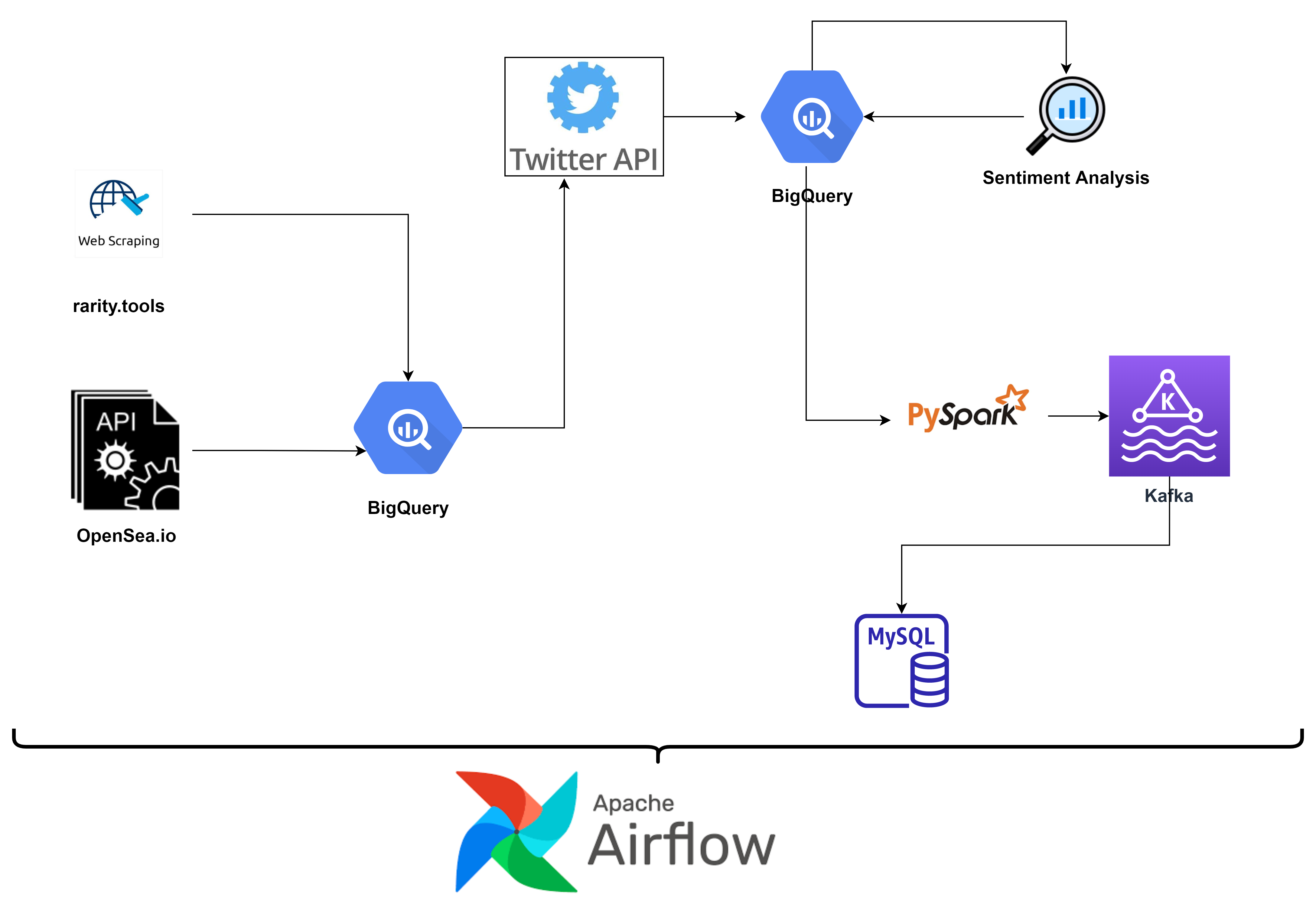
NFT is a dynamically growing market with hundreds of new projects that announced every day. In same time it has big investment potential.

Analysis of NFT projects that based on project details and social media parameters can predict the value of the item and allow collectors, investors and other people to choose their purchase, with consideration of future value.

It helps to be more focused on specific projects, and not to waste time on investigation of projects, with less potential value.

In this project we considered how such analysis can be implemented.

**Project Architecture Diagram:**



**Tools & Technologies:**

**Python libraries**

*Google-cloud-bigquery*

API Client for working with Google bigquery

*Beautifulsoup4*

Package for parsing HTML data

*Selenium*

UI automation for web scrapping

*Twython*

Python wrapper for Twitter API, used to get the data from twitter  
*Requests*

Library for HTTP request performed by OpenSea API  
*Uuid*  
 Library for creating immutable UUID objects (unique key generator), used to generate Id for our data tables

*Datetime*

Library that supplies classes for manipulating dates and times.

**Shell:**

Preparing environment before project run: Installing Python libraries

GIT Cloning latest version of Python project

Chrome browser Installation

Chrome driver Installation

**BigQuery:**

BigQuery is a fully managed enterprise data warehouse that helps you manage and analyze your data with built-in features like machine learning, geospatial analysis, and business intelligence.

Used as main storage system of projects’ data.

Main benefits: native GCP integration, Serverless, scalable & free with trial credits

**MySQL:**

An [open-source](https://en.wikipedia.org/wiki/Open-source_software) [relational database management system](https://en.wikipedia.org/wiki/Relational_database_management_system) (RDBMS)

Used to store the final results of our rated NFTs

**Web Scrapping:**

An approach for collecting data from a website, when no API is available.

Used to collect data on latest and upcoming projects from Rarity.tools URL.

**REST API:**

A REST API (also known as RESTful API) is an application programming interface

Used for collecting data from NFT’s trading platform OpenSea.com

**Twitter API:**

The Twitter API enables programmatic access to Twitter in unique and advanced ways

Used to get accounts, twits and statuses data, from twitter social network by API calls.

**Sentiment Analysis:**

Google Sentiment Analysis performs sentiment analysis on text. Sentiment analysis attempts to determine the overall attitude (positive or negative) and is represented by numerical score and magnitude values.

Used for calculation of twits' text sentiment and assignment of relevant score.

**Spark:**

Apache Spark is a multi-language engine for executing data engineering, data science, and machine learning on single-node machines or clusters.

Used for getting data from Bigquery tables and data processing, including the final rating calculation.

**Kafka:**

Apache Kafka is an open-source distributed event streaming platform used for high-performance data pipelines, streaming analytics, data integration, and mission-critical applications.

Used as a hose to transfer the final results to MySQL DB.

**Apache Airflow:**

Apache Airflow is an open-source workflow management platform for data engineering pipelines.

Used to orchestrate and coordinate the tasks in our pipeline.

