**Project 2, Group 17’s Proposal**

Jonathan Michel, Haslett Fernandes, Justin Ho

The Trilogy

**December 8th, 2022**

# Project Description

Looking at popular movie review sites such as Rotten Tomato, IMDb, and Metacritic, we have noticed varying review scores amongst the three for the same movie. For example, the movie “Pain and Glory” received a 7.5 from IMDb, an 87 from Metacritic, and a 96 from Rotten Tomatoes. Each site has their own user base with their own grading scale to review movies whether it be from critics or the general public. Therefore, for this project we wanted to look at the film industry and extract the data of movie review scores across the three sites. The data sources we will be extracting from are the OMDB API and the Metacritic CSV file found on Kaggle. Once the data has been loaded in, we will be combining the data from the OMDB api into a dataframe with the data from the Metacritic CSV. Finally, once the data has been cleaned and formatted, we will be loading it into a SQL relational database.

# Data Sources:

* [OMDB Api](https://www.omdbapi.com/)
* [Metacritic CSV](https://www.kaggle.com/datasets/miazhx/metacritic-movie-reviews?select=metacritic_movies.csv)