

Data Engineering Projekt 2

Ilef Kalboussi (81774)
Nina Scheiffele (87488)
Kevin Garrison (85826)

28 November 2024

1 Beschreibung der Szenarioerweiterung

Dieses Datenbankmodell unterstützt eine Streaming-Anwendung (Netflix) für Medien, die es den Nutzern ermöglicht, Watchlists zu erstellen und Medieninhalte zu bewerten, mit einer strukturierten Verwaltung von Abonnements und Medientypen. Im weiteren wurde die Ausgabe und das Einlesen von JSON Dateien hinzugefügt.

Eingabedateien erstellen

- **Einfaches Beispiel für eine einzelne Entität:** Eine Datei wurde erstellt, die einen `OtherUser` als einzelne Entität enthält und direkt in die Datenbank eingelesen wird.
- **Komplexeres Beispiel mit Extraktion/Parsing:** Eine JSON-Datei wurde vorbereitet, die einen `Director` und zwei `Actors` enthält. Der Inhalt dieser Datei wird geparst, um unterschiedliche Entitäten sowie mehrere Instanzen einer Entität in die Datenbank zu importieren.
- **Mehrere Datensätze in einer Datei:** Der Aufgabenpunkt wurde mit der vorangehenden Lösung bereits abgedeckt, indem zwei `Actors` aus der JSON-Datei eingelesen werden.

Ausgabedateien aus der Datenbank generieren

- **Einzelne Entitäten exportieren:** Zwei separate Dateien wurden erstellt, die jeweils eine Instanz der Entitäten `Review` und `Watchlist` enthalten.
- **Abfrageergebnisse exportieren:** Eine weitere Datei wurde generiert, die die Ergebnisse folgender Abfragen enthalten:
 - `count_reviews_per_media`: Anzahl der Rezensionen pro Medium.
 - `average_rating_by_genre`: Durchschnittsbewertung pro Genre.
 - `count_episodes_per_series`: Anzahl der Episoden pro Serie.