

# Recommender Systems

Prof. Dietmar Jannach  
Alpen-Adria-Universität Klagenfurt

## Assignment 7: Project “Similar Item Recommendations”

The goal of the concluding project is to experiment with different strategies to generate recommendations of similar items. The movie domain can be used for this experiment.

Generally, similar item recommendations can be found on many websites, including video streaming sites. It is, however, not always immediately clear when we should consider two movies to be similar. It could be because they have the same actors, the same director, similar plot descriptions, the same genre, or even a similar movie cover. A recent research work on the topic can be found here:

[https://web-ainf.aau.at/pub/jannach/files/Journal\\_UMUAI\\_2019-2.pdf](https://web-ainf.aau.at/pub/jannach/files/Journal_UMUAI_2019-2.pdf)

The first task in this project is to develop a number of (at least 5) functions in Python that, given a reference movie ID, return a ranked list of the top-5 most similar items. Each function has to implement a different strategy.

For the experiment, you can use the MovieLens 20M (or a smaller) dataset, which also contains some content information, see: <https://grouplens.org/datasets/movielens/>

In addition, a number of additional movie features can be found here:

<https://drive.google.com/file/d/1je77e0Lq8naVUsjoOzk5Rul2H3ceHISz/view> )(500 MB)

This zip file contains a number of JSON-files, each file containing additional data for a given movie (but no ratings). The MovieLens dataset can be combined with this dataset based on the field “movielensID”.

In this dataset, many of the URLs pointing to the movie poster images are outdated. An alternative dataset with posters for the movies can be found here:

<https://www.kaggle.com/ghrzarea/movielens-20m-posters-for-machine-learning>

Next, implement a user interface, where the reference movie can be searched by title. The user can then select one of the returned search results and the system then presents five lists, each one containing similar movie recommendations. Each list is based on one of the implemented similarity functions.

The user interface has to be web-based (e.g., using the Django framework) and it has to be designed in a way that it is reasonably easy to use in terms of searching, movie selection, and the analysis of the results.

The final task lies in the evaluation of the different strategies. Try out at least 20 movies from different genres and manually inspect the different result lists. Write a short report that summarizes the observations. Possible contents could be: Which recommendation method led to the best results? Which method led to surprises? Which method led to poor or unexpected results etc.

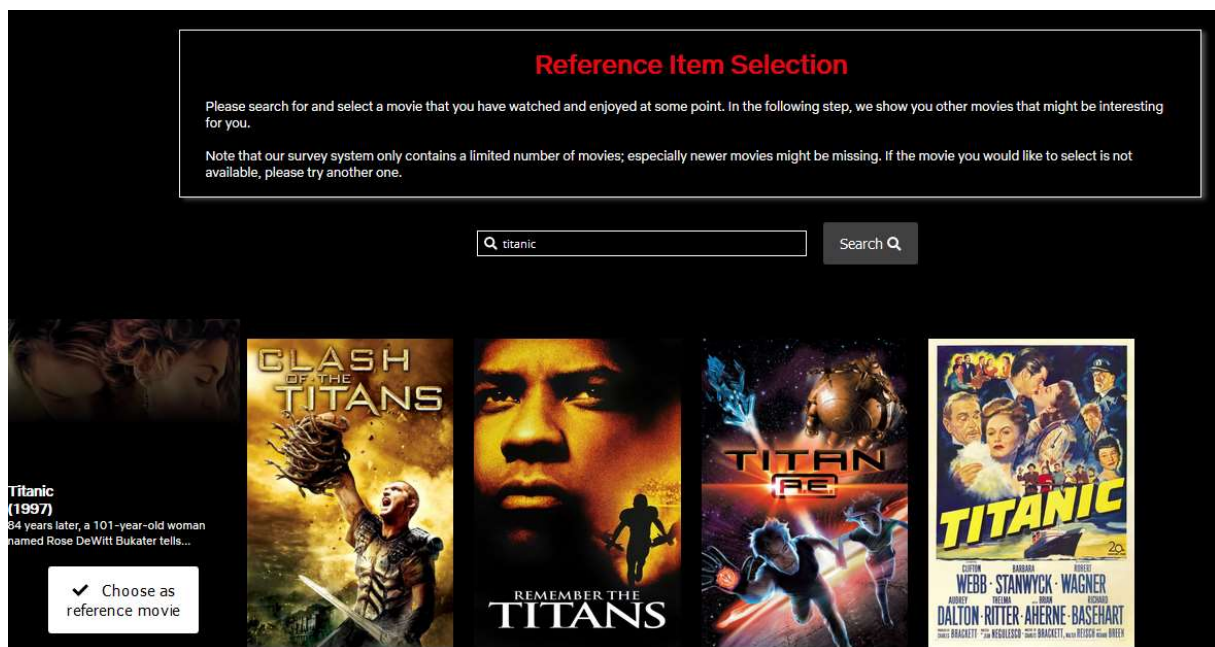
## Project planning / important dates:

- **Intermediate meeting:** Each project team presents an update on the current state of their design and implementation of the project. Presentation time per team is 20 minutes.
- **Final meeting:** Each project team presents their final software solution and the outcomes of their analyses. Presentation time per team is 30 minutes max.
- **Final deadline:** Deadline for submitting the project code including light documentation and the report by email to dietmar.jannach@aau.at

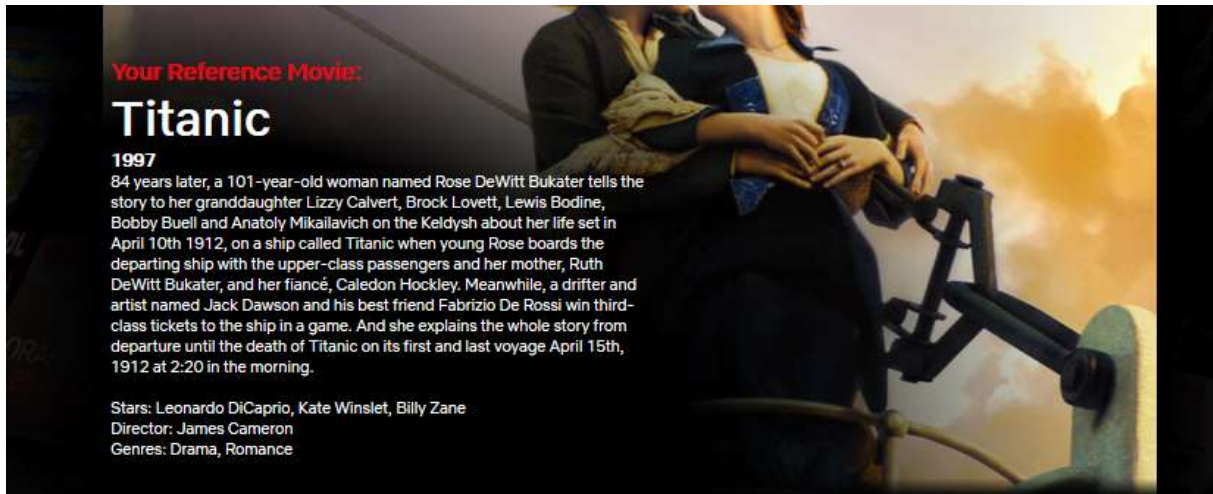
## Screen designs:

Find below examples of an interface design for an ongoing user study that can serve as inspiration.

### 1) Movie selection



- 2) Result presentation: Here it is just one list of similar movies – in your project, you will have multiple rows with different labels for the lists.



**Your Reference Movie:**

# Titanic

**1997**

84 years later, a 101-year-old woman named Rose DeWitt Bukater tells the story to her granddaughter Lizzy Calvert, Brock Lovett, Lewis Bodine, Bobby Buell and Anatoly Mikailavich on the Keldysh about her life set in April 10th 1912, on a ship called Titanic when young Rose boards the departing ship with the upper-class passengers and her mother, Ruth DeWitt Bukater, and her fiancé, Caledon Hockley. Meanwhile, a drifter and artist named Jack Dawson and his best friend Fabrizio De Rossi win third-class tickets to the ship in a game. And she explains the whole story from departure until the death of Titanic on its first and last voyage April 15th, 1912 at 2:20 in the morning.

Stars: Leonardo DiCaprio, Kate Winslet, Billy Zane  
Director: James Cameron  
Genres: Drama, Romance

## SIMILAR MOVIES

