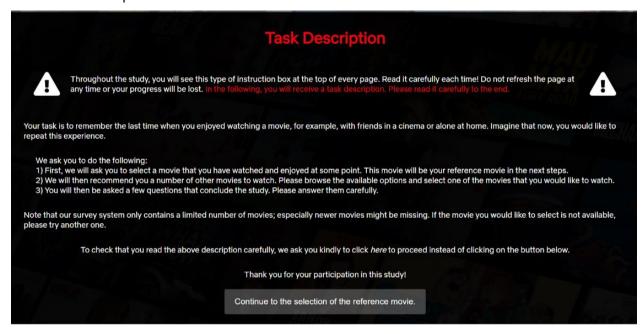
Additional material for UMAP 2021 paper submission

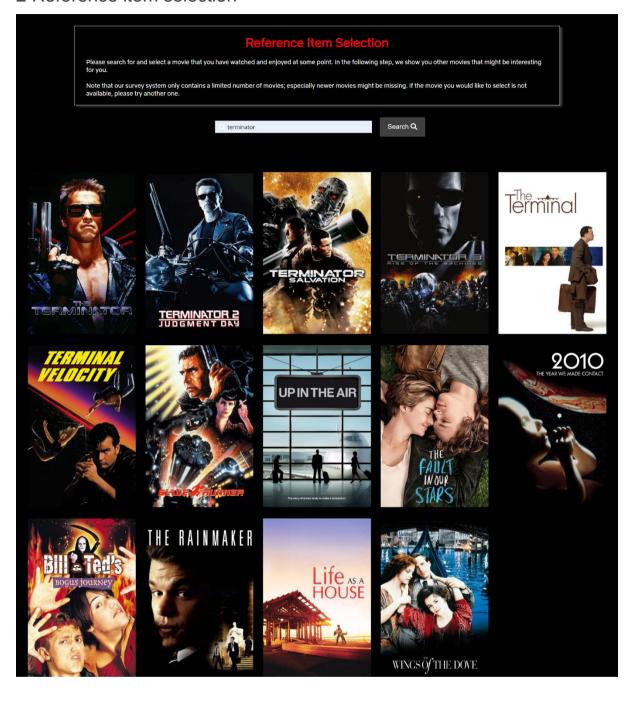
"Exploring Multi-List User Interfaces for Similar-Item Recommendations"

Online Study: Screen Captures

1 Task description

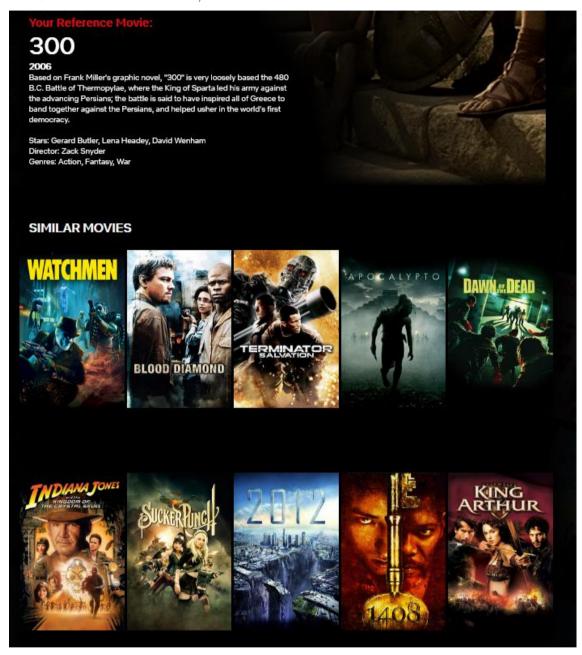


2 Reference item selection

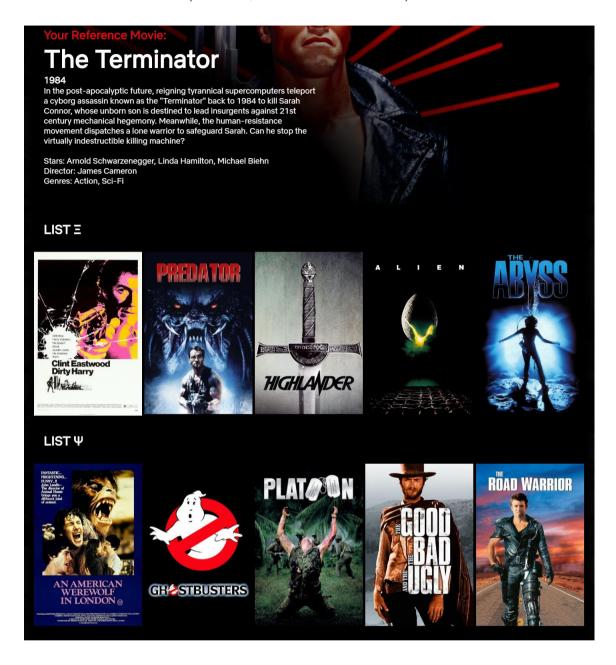


3 Presentation of items

3.1 Condition with one row, line breaks and one list label

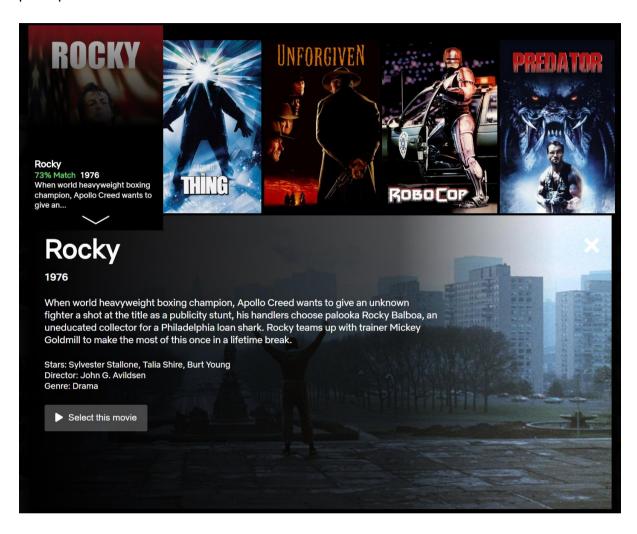


3.2 Condition with multiple rows, line breaks and multiple list labels



4 Detail view of a movie

When clicking on a movie cover, more information was displayed along with a button where participants could make their selection.



Post-Task Questionnaire

Questionaire (Statement to rate)	Key in Output Data	Possible values
It was difficult to find similar movies.	q_difficulty1	Likert scale (1-7)
All of the presented movies were similar to the reference movie.	q_similarity	Likert scale (1-7)
Searching for similar movies took a lot of time.	q_time_effort	Likert scale (1-7)
Deciding between the large number of options was difficult.	q_nb_alternatives	Likert scale (1-7)
The arrangement of the movies was well organized.	q_order	Likert scale (1-7)
It was difficult to find a suitable movie.	q_difficulty2	Likert scale (1-7)
I found more than one suitable movie.	q_more_than_once	Likert scale (1-7)
The proposed movies were diverse.	q_diversity	Likert scale (1-7)
The number of options was sufficient.	q_sufficient	Likert scale (1-7)
I was satisfied with the number of options.	q_satisfaction_with_options	Likert scale (1-7)
I am familiar with the movie that I chose.	q_decision_familiarity	Likert scale (1-7)
I am satisfied with the movie that I chose.	q_decision_satisfaction	Likert scale (1-7)
I was able to discover new movies.	q_novelty	Likert scale (1-7)
The application was easy to use.	q_ease_of_use	Likert scale (1-7)
Searching for a suitable movie was intuitive.	q_intuitive	Likert scale (1-7)
The situation described in the study was realistic	q_realistic1	Likert scale (1-7)
I had no difficulty imagining myself in the situation described in the study.	q_realistic2	Likert scale (1-7)

Personal information	Key in Output Data	Possible values
Your age:	q_age	Textfield
Your gender:	q_gender	Dropdown (Male, Female, Other)
Which of the following statements best describes your use of online movie services (e.g, Netflix, IMDB, etc.)?	q_streaming_consumption	Dropdown (Daily, Once a week, Once a month, Every 3 month, Hardly use them)
On average, how many days per week do you watch a movie?	q_movie_consumption	Dropdown (0 to 7)
General comments	q_comment	Textfield

Objective Measurements

Documentation of	
Data	Description
session_id	Unique identification of the session
start_time	Timestamp the participant started the survey
condition	The condition group the participant was assigned
ref_item	The movie the participant chose as a reference
choice	The recommended movie the participant chose
choice_list_index	The index of the list in which the recommended movie is that the participant selected
choice_item_index	The index of the recommended movie participant selected
choice_list_name	In condition #4, #5 and #6 this attribute is used to group different lists
choice_time	How long did the participant take to make a choice
search_time	How long did the participant take to search
failed_first_check	Did the participant fail the first check?
failed_second_check	Did the participant fail the second check?
intro_time	How long did the participant take for the intro
survey_time	How long did the participant take for the survey
recommended_movies	What movies have been recommended to the participant
nb_hovered	Over how many movies did the user hover with the mouse
nb_unique_hovered	Over how many unique movies did the user hover with the mouse
overall time (derived)	How long did the participant take for the complete study