Utrecht University

Department of Information And Computing Sciences

Game And Media Technology - Master Thesis

Comparing Visual Exploration and Standard Interfaces for Movie Browsing

Author:
Kylian Kuijer

Supervisor:
Dr. Wolfgang Hürst

Second Examiner:
Prof. Lynda Hardman

ICA-4055195

July 17, 2018

Universiteit Utrecht

Summary

As the number of people using movie browsing applications increase annually, so does research into the demands of these users. In this field, state-of-the-art interfaces frequently emphasize on recommendations to provide users with content to watch. Recommender systems have their own weaknesses, however: users may become trapped in a bubble, where they have trouble finding items that the system does not recommend to them. As such, users may have trouble stepping outside their comfort zone. Recommender systems may also fail to accommodate for temporary changes in mood, instead looking at general behavioral patterns. Many also require established data for both users and the items present in the database, as can be seen in collaborative-filtering recommender systems. These issues may cause state-of-the-art interfaces to fail in satisfying the needs of exploratory search. In this thesis, we focus on a recently introduced interaction concept, called the *MovieWall*, which addresses these issues. It does so by emphasizing on exploration and discovery, letting users visually navigate a movie collection presented as a large grid of movie posters, specifically made for mobile, touch-based devices.

Initially, results from a related user study were positive, indicating that the *MovieWall* could be promising as a complementary feature to existing approaches. Yet, no direct comparison was made with the state-of-the-art. As such, we performed a comparative study to identify the strengths and weaknesses of the *MovieWall* with regard to existing interfaces. Results of this study indicate that the current implementation of the *MovieWall* is lacking in various aspects. Users perceived a mock-up of the state-of-the-art interface, called the *Standard* interface, to be more practical and better suited to their needs. In an interview, participants specifically mentioned the perceived lack of structure and the lack of guidance, which caused them to get lost easily and impacted their experience negatively. Many stated they enjoyed the aspects of discovery and found the interface fun and exciting. As such, we believe that the concept of the *MovieWall* may still find use as a complementary feature once the aforementioned issues are resolved in future implementations.

In short, our user study showed that the current *MovieWall* implementation is inadequate as a complement to or replacement of state-of-the-art interfaces. However, we also identified areas of improvements which, once solved, could still make the *MovieWall* a valuable complement to existing interfaces.

Acknowledgements

I would personally thank the following people:

- My supervisor, dr. Wolfgang Hürst, for his guidance and feedback during this thesis project.
- My second supervisor, Prof. Lynda Hardman, for taking her time to evaluate this thesis project.
- Friends and family, who gave me very valuable feedback from a non-academic point of view and frequently assisted me in keeping focus on what is most important when researching usability and user experience.
- Participants of the user study, who helped me gather valuable data and brought to light important fact, as well as provided many possibilities for future research.

Without your support, this thesis project would not exist today.

Contents

Introduction

In this thesis, we compared the *Standard* interface, a mock-up of the state-of-the-art interface, with the *MovieWall*, an interface concept that utilizes visual exploration of movie collections on mobile tablet devices to encourage discovery and exploratory browsing. The following is contained within the document:

- A scientific paper with the major scientific results of the thesis.
- An annotated appendix containing all material relevant to the project, including a more in-depth discussion of some aspects from the scientific paper plus further details, e.g. about related implementations. In particular:
  - A literature review on the progression and comparison of browsing systems, with emphasis on query-based browsing, recommender systems and exploratory search The results of this review formed the bases for the research of this thesis.
  - Material containing an in-depth explanation of the complete implementation used for the comparative user study. This includes a flow model of the *Standard* interface, as well as future research opportunities regarding the *MovieWall* and its current limitations.
  - An extensive description of the comparative user study, including methodology, additional notes, design decisions and all data gathered and compiled.
- A final conclusion to this document and pointers towards future work based on related works, the literature review and the conclusion of the scientific paper.

Further deliverables of this thesis, which are not included in this document are:

- The source code, executables and tools used for the comparative study, including explanation for future users.
- All data gathered during the user study, which consists of:
  - Raw log files.

  - All results from the online form.
  - Interview notes.
  - A compilation of statements made by users within the interview, including references to the interview notes for each statement.

- Dissemination material targeted at a wider audience, showing both interfaces and key results of the comparative user study.

Scientific Paper

The following pages contain the scientific paper, which summarizes the main scientific results of the thesis. In the paper, the *Standard* interface and the *MovieWall* interface are compared and the outcome of the comparison is discussed in detail, with pointers towards future concepts and improvements.

Comparing a Visual Exploration interface for Large Movie Collections with Standard Interfaces

Kylian Kuijer
Utrecht University
Utrecht, The Netherlands
k.kuijer@students.uu.nl

July 17, 2018

Abstract

As movie streaming services become ever more popular, demand rises for interfaces that give users a great experience while browsing. A concept that emphasizes on this experience was recently realized as the *MovieWall*, an approach where users explore movie collections with little restriction [8]. The *MovieWall* lets users visually explore content, without restricting freedom, allowing them to satisfy exploratory search needs that may not be satisfied by state-of-the-art interfaces. While initial results appeared positive, no direct comparison was made between the two. A comparative user study using a mock-up implementation of the state-of-the-art, the *Standard* interface, shows the *MovieWall* implementation to be inadequate, with few users preferring it over its competitor. Users largely favored the *Standard* interface in terms of usability and user experience. Statements from subject interviews show that the *MovieWall* was interesting and fun as a concept, with special mention to its emphasis on exploration. Yet, participants considered the current implementation to be lacking and cumbersome for various reasons, including; a lack of guidance, a lack of proper structure and the inability to personalize the interface to any degree. Interestingly, subjects did not think the database behind each system was different, though the *Standard* interface gave them an easier time finding movies of interest. Both interfaces seem to scale to multiple screen sizes without major issues, even hinting that both concepts could work on different devices altogether. Further work on improved implementations of the*MovieWall* concept should be done before recommending it as a complementary function, Future implementations must improve usability and provide more guidance without limiting the freedom of the user. Additionally, new iterations of the *MovieWall* should have a better underlying structure and add personalization options based on user input. We believe the *MovieWall* concept to be a suitable complement once these improvements are made.

1 Introduction

Online video services provide various features to browse their collection of movies. These range from a simple search bar granting users the opportunity to query the database, to the recommender system showing users related items depending on various criteria. As collections become ever larger and more diverse, so too does the need for systems to provide content which is interesting to the user. Their needs may change drastically, however: while they may enjoy the content provided on one day, they could prefer something drastically different on the next day.

State-of-the-art movie-browsing interfaces such as Netflix generally consist of a search function, followed by recommendations/ These are presented in lists with items horizontally-oriented, and each list containing a header that describes the content for that list. Features may be present that alter the content of each list. The metadata used to generate each list can vary greatly; from tags frequently used for these items (like genres, e.g. "Action" and "Adventure") to popularity and more.

Such interfaces have various weaknesses. The user is fre-

quently forced to only see a subset of the collection. In the event an interface presents recommendations of limited interest, the user may not find anything of value unless they either navigate through the system manually or discover the query to enter into the search function that will bring them to their destination. The former is generally considered cumbersome, as it requires users to navigate several pages without clear knowledge of how the system is guiding them through the content, potentially never reaching their destination. The latter requires them to iteratively improve their query to find the content they desire, which can be equally cumbersome and may be impossible. Furthermore, such an interface may frequently hamper a user's freedom in order to provide guidance. This may not be desired by users when they wish to explore the movie collection at their leisure. In such contexts, the restriction in freedom may be seen as a significant detriment. Finally, recommendations require various data to be effective, which could be private and it requires users to have been using the service for some significant length of time (known as the "cold start" problem, specifically a "cold user" [1]).

M. Nefkens introduced an approach entitled the *MovieWall* [8] that emphasizes user freedom and minimum restriction when exploring movie databases. A prototype of this interaction design, which is optimized for touch interaction, has been implemented and evaluated on a state-of-the-art tablet device. The application consists of an interactive interface where movie posters are arranged in a tile-based lay-out by some form of clustering. Users navigate this lay-out by performing pinching and dragging gestures. Tapping on individual movie posters provides the user with various details and presents them with additional actions to perform, such as adding the selected movie to a watch list. The interface largely functions as a single-page application. Users can also apply various filters to highlight items of interest.

Subjects of a user study involving this prototype reacted positively; twenty-four out of thirty participants said they enjoyed the application and twenty-six said they would use a similar interface in a real movie-browsing application. These findings suggested that the concept had confirmed the usefulness of the design as a complementary approach to state-of-the-art interfaces. Some evidence even indicates that it might be a suitable replacement. Yet, this experiment only studied several implementations of the *MovieWall*, but did not compare it with standard solutions. The purpose of this research is thus to verify these indications, that is, to answer the following question: Is the concept of the *MovieWall* a complement to existing state-of-the-art approaches and could it even replace them completely?

To answer this question, we conducted a comparative user study between a mock-up of a state-of-the-art interface, subsequently referred to as the *Standard* interface, and a *MovieWall* implementation. Our work makes the following contributions:

- We provide a comparison between the *Standard* interface and the *MovieWall* with respect to usability and user experience:
  - Differences in usability of the interfaces addressing total and individual aspects.
  - Differences in the experience with respect to their visual representation and the achieved browsing behavior.
- We evaluate potential differences in the perception of the collection:
  - The composition of the database as a whole, for example, if it is perceived as different.
  - In which way users prefer items in the collection to be provided and which interface they believe presents them with more items interesting to them.
- We investigate whether both interfaces are portable to different tablet sizes without any complications.
- Based on these observations:
  - We prove or disprove the suitability of the *MovieWall* as a concept for engaging and satisfying exploration of movie databases.
  - We identify pros and cons, provide guidelines for successful usage based on the potential advantages and pinpoint necessary alleys for future research to cope with the observed disadvantages.

The paper is structured as follows: Section 2 highlights related work in the fields of recommender systems, exploratory search and recent applications regarding movie and video browsing. In Section 3, we describe the comparative user study, including details about the evaluation

Figure 1: The *MovieWall* interface. On the left side, one can see the interface completely zoomed out, which also serves as the starting point. The headers show the centers of each cluster, as well as the genres used. On the right, a view of the *MovieWall* zoomed-in, with the details panel open. Headers disappear as the user zooms in. All movie posters were downloaded from www.themoviedb.org and used under the fair usage policy.

and experiment design. Section 4 summarizes the results, discusses and analyzes them. Finally, we summarize our conclusions in Section 5 and address important aspects for future work.

2 Related Work

Acknowledging the wide spread use and potential of recommender systems, answers to the weaknesses mentioned before are investigated extensively. In the area of user privacy, McSherry et al. used the Netflix Prize data set and adapted its leading algorithms to provide differential privacy while losing minimal accuracy [7]. However, such systems still require significant user input, which may not be available. A different implementation resolving the issue of cold start items was made by Wei et al, which combines time-aware collaborative filtering and deep learning to predict ratings for new items [10]. Their implementation was also tested on the Netflix dataset with large improvements. While it is clear that the weaknesses of recommender systems are increasingly mitigated, they are still present and may always be to some degree. As such, alternative interfaces to state-of-the-art systems which do not rely on private user data or requires users to have used the system extensively are worth exploring. For our research, both the *Standard* interface and the *MovieWall* are such interfaces, as they do not use any data from participants.

Work in the field of exploratory search, which is an important part of the *MovieWall*, has been on the rise. The importance of exploratory search has been highlighted by several writers, including the works of Pang et al. [9] and Marchionni [6]. Users have difficulties finding what they desire when they are unable to describe it in a way that can be queried. Accepted definitions and characteristics of exploratory search are very open [11]. Under these definitions, interfaces that emphasize on recommender systems may answer any exploratory search needs of the user. Yet, given the weaknesses described before, providing alternatives to exploratory search is important. As such, the strengths and weaknesses of state-of-the-art interfaces and new concepts like the *MovieWall*, should be investigated thoroughly.

A solution which also has the user explore by visualizations was presented in the work by Carvalhal et al. [3]. In their interface, metadata such as genres, actors, directors and companies is visualized rather than presented as text. These visualizations can be tapped add the metadata to a query, updating the movies shown based on this query. By progressively adding or removing metadata, users evolve this query, navigating through the collection. In a direct comparison, results show that users enjoyed this method of navigating and visualization more

Figure 2: The *Standard* interface. The home screen is shown on the left side, where the user starts. The right side shows what the user tapping on a movie, which opens the details panel. All movie posters were downloaded from www.themoviedb.org and used under the fair usage policy.

than traditional state-of-the-art interfaces, without influencing usability and usefulness. One point of concern may be a bias in results due to novelty, which could fade in the long-term. However, the results appear to be promising and could be investigated further. While their work still relies on queries and filtering, it provides a different venue to leisurely browsing in a fun way.

3 User Study Design

To answer the question of how the *MovieWall* concept compares to the *Standard* interface, we present a comparative user study. We start by further specifying the objectives into sub-questions of the experiment in Section 3.1. In Section 3.2, we describe the implementation and motivation for related design decisions. The study design is addressed in Section 3.3.

3.1 Objectives

As stated in the introduction, our major aim is to verify if the *MovieWall* interface is an appropriate complement or even replacement of a state-of-the-art interface implementation. We can answer these questions positively, if it is either better in niche situations, preferential for a significant subset of users, or overall superior to the *Standard* interface.

The objective of the first sub-question is to compare the usability of both approaches. We are interested in general usability as well as identifying sub-aspects that could pinpoint scenarios or contexts where one approach might be preferable over the other, or where the *MovieWall* implementation is still lacking. Furthermore, we are interested in the reasons for any shortcomings. This helps future research solve existing issues and improve the *MovieWall*. As we discussed above, we expect the *MovieWall* approach to be particularly helpful in scenarios for exploratory search, especially when the search goal is vague, badly specified, or non-existent. We therefore phrase this sub-question as:

SQ1: How do the two interfaces compare in terms of usability for a general browsing task with vague search goals?

One major advantage of the *MovieWall* is its ability to show a much larger subset of movies at any point in time than any of the standard implementations [8]. Given this fact, we assume that it will be easier for people to get a better overview of the database more quickly, as they obtain information on the movie collection at a much faster pace. Furthermore, being able to access movie data more rapidly may decrease the likelihood of missing relevant data completely, as users are more likely to encounter anything of interest when the number of movies they come

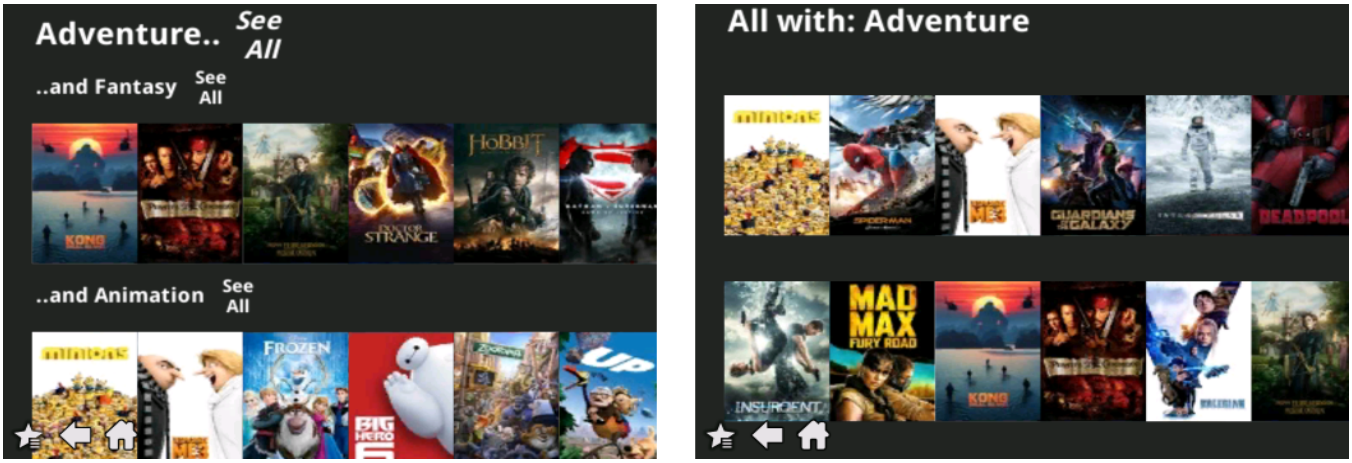

Figure 3: The view on the left shows the "See More" function in action. The view on the right showcases the "See All" function. All movie posters were downloaded from www.themoviedb.org and used under the fair usage policy.

across increases. Because evaluating this aspect is difficult, we verify it in a different way. We let users explore the very same database with both interfaces, but without mentioning the equality of the data. We then argue that if one of the databases is perceived as âĂIJbetterâĂİ, that is, it users conclude that it contains more movies that are relevant to them, the interface used for this case provides a faster and easier access to the relevant files. Thus, we phrase the second sub-question as:

SQ2: How does the interface influence the perceived quality of the explored database?

Finally, the *Standard* interface is a proven concept that can be applied to multiple devices of various screen sizes. As a competitor or complementary function, we expect the *MovieWall* to work on tablets of various sizes as well, without major constraints or issues. This also provides a frame of reference when theorizing how the *MovieWall* interface would perform on different types of devices, such as smartphones or smart TVs. As such, we formulate subquestion SQ3 as follows:

SQ3: Does screen size have a significant influence on any of the aforementioned aspects?

3.2 Implementation

The *Standard* interface was embedded in the same application as the *MovieWall*, which was developed using Unity3D [8]. The application was made for Android and designed for tablet use. The application supports two resolutions; 4:3 using the 9-inch HTC Nexus 9 as reference, and 16:10 using the 7-inch Asus Google Nexus 7 as reference. These two tablets form the basis of comparison for SQ3. The database used by both interfaces consisted of 2160 movies. The data of all movies used were provided by The Movie Database (TMDb). Before moving to either interface, the system requests users to submit genres they prefer. These choices are stored and used for recommendations by the *Standard* interface. Users may only select genres which both interfaces use for their primary representations. The genres used are the same as in [8], and can be seen as labels in Figure 1, on the left. Each interface utilizes the same underlying database, with the *MovieWall* displaying all movies within the collection at once, while the *Standard* interface shows subsets based on filter criteria. Users can tap individual movie posters in either interface, which provides access to the movie details panel. Here, the user can find additional information about the movie, as well as add the movie to the watch list. In both interfaces, metadata are shown underlined, which can also be tapped, though functionality differs for each interface. The differences are discussed below. The*MovieWall* implementation used in the experiment is based on [8]. We evaluated various extensions

S1	I think that I would like to use this system frequently.
S2	I found the system unnecessarily complex.
S3	I thought the system was easy to use.
S4	I think that I would need the support of a technical person to be able to use this system.
S5	I found the various functions in this system were well integrated.
S6	I thought there was too much inconsistency in this system.
S7	I would imagine that most people would learn to use this system very quickly.
S8	I found the system very cumbersome to use.
S9	I felt very confident using the system.
S10	I needed to learn a lot of things before I could get going with this system.
A1	Which service did you prefer overall?
A2	Which visual presentation did you prefer?
A3	Which service gave you a better experience?
A4	Please rate which service had a movie collection better suited to you.
A5	Please rate which service made it easier to find movies interesting to you.

Table 1: Statements based on SUS questionnaire asked after completing the task for one interface (prefix S) and additional comparative statements asked after completion the task with both interfaces (prefix A). Scores use a 7-point Likert-scale. For SUS statements, 1 equals "heavily disagree" and 7 equals "heavily agree". For additional statements, 1 favors the *Standard* interface and 7 favors the MovieWall.

and modifications in order to improve it for our test. In the end, we decided to use the very same functionality, as we believe a better comparison can be made with previous research if a tested implementation of the *MovieWall* is used. One key difference in our implementation is to start the *MovieWall* at maximum zoom level, as early evaluations indicated that users preferred the labels giving them a general overview of the lay-out. When selecting the *Standard* interface, users start at the "Home screen". Here, horizontal lists are generated for each genre shown in Figure 1, including a descriptive header. Items inside each list are sorted by popularity. Genres that the user selected as their preferences are shown first, with the other genres being sorted to the bottom. Horizontal lists can be dragged horizontally to scroll through up to thirty (most popular) entries, while dragging vertically anywhere on the screen allows the user to scroll the lists themselves. With the exception of one view, all views follow this same pattern, only changing in the content of each list. Figure 2 shows this in more detail. From here, users have a few options. Opening the details panel by tapping on any of the movies also shows the "More Like This" button at the bottom of the panel, as seen in Figure 2 (right side). This feature allows them to load new lists, using the selected movie as reference: the lists will be filled with movies that have the same genres, actors, directors and companies as that movie. Furthermore, users can tap on "See More" near each header, which then presents the user with lists that combine the genre of the header with every other genre used for clustering. An example of these features can be seen in Figure 3. To allow the user access to the entirety of the database in a few taps, we introduced the "See All" view. Using the "More Like This" and "See More functions changes the "See More" text near the header to "See All". Additionally, using the "See More" function generates a "See All" button next to the header, at the top of the screen. When tapping "See All", a slightly different view is loaded; a filter is applied to the entire database depending on the header that was tapped, e.g. tapping on "See All" next to the header with "Adventure" will filter on Adventure movies. All movies are then similarly sorted by popularity and presented in a large list which the user can only drag vertically, though movies can still be tapped and viewed. This view can be accessed by pressing the filters (underlined words) in the details panel as well. This function is shown in Figure 3.

The *Standard* interface includes two more functions, accessible from any view. The first is a "Home" button, which returns the user to the home screen. The second is a "Back" button, which restores the previous view when pressed, allowing users to backtrack conveniently.

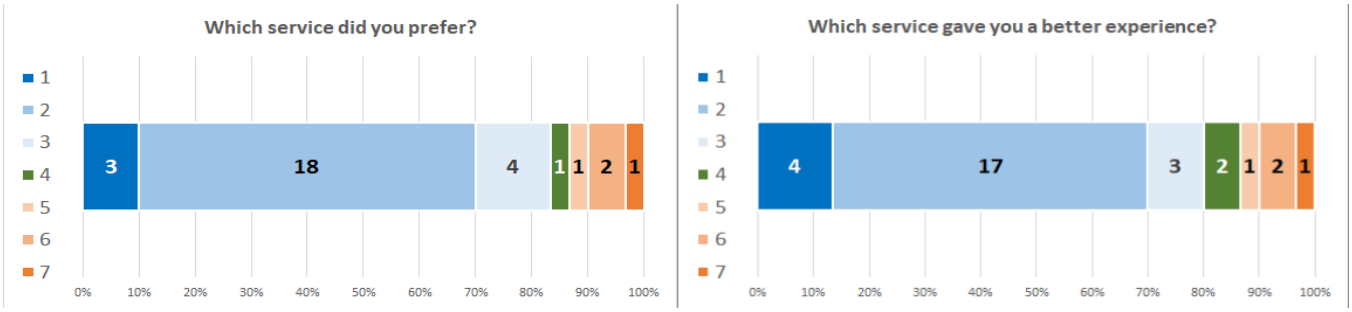

Figure 4: Stacked bar figures for the preferred service ratings and preferred experience ratings. Blue segments show in favor of the *Standard* interface, orange segments in favor of the *MovieWall* interface and green as neutral.

3.3 Procedure

Experiments took place in a controlled environment with an observing party. Participation was entirely voluntary and not reimbursed. Subjects gave their consent prior to the experiment by filling their email address into an online form, receiving a copy of their answers at completion. Addresses were not tracked and removed after participation. Users also filled in general background questions regarding their use of movie browsing and video-on-demand services. Participants were divided equally among the HTC Nexus 9 and the Asus Google Nexus 7, and tested the application on their assigned device. All subjects were given a task to perform on each interface. This task was designed to stimulate users to explore and browse both interfaces, simulating exploratory search. As such, the task was open-ended, encouraging users to base their preference on the user experience and database perception, in parallel with our sub-questions. To further emphasize this, participants were given little time to complete it. We described the task as follows:

Assume you have won a free one-year subscription to an online video service, and you can choose between two offerings, each with their own video-on-demand service. Go ahead and spend about four minutes checking out each of them, and make a decision which of the two you would like to subscribe to. Users tested the *Standard* interface first, followed by the *MovieWall*, so all participants had a grasp of the state-of-the-art before comparing. After completing the task for an interface, subjects answered a SUS questionnaire [2] for that interface before moving on. After completing the task on both interfaces, participants gave answers to a series of questions, primarily consisting of comparative ratings. These statements can be seen in Table 1. Answers to the SUS questionnaire, as well as statements A1-A3, are used to measure the user experience and usability of each interface (SQ1). Statements A4 and A5 provide insight to perception of the database for each subject (SQ2). Both the SUS questionnaire and the comparative questions were rated on a seven-point Likert-scale for consistency. SUS-scores were normalized to range from 0 to 100 (five-point Likert-scale) afterwards for evaluation using the SUS standard. For comparative questions, a score of 1 heavily favored the *Standard* interface, while a score of 7 heavily favored the *MovieWall*, with 4 being the neutral score. We compared the ratings given for the Nexus 7 and Nexus 9 groups to answer whether the screen size had any influence on user experience or database perception (SQ3). Afterwards, users would participate in a short interview, presenting qualitative statements for each interface. Along with observations, these were used to provide additional information regarding the user experience and their perception of the movie collection of each interface. These statements were combined and tallied to find phrases that were mentioned frequently. All actions performed by the users were timestamped and logged. Each action was totaled for each user, so we could look for features that remained unused a participant, i.e. the action being performed zero times. Every feature had a unique action that identified its execution.

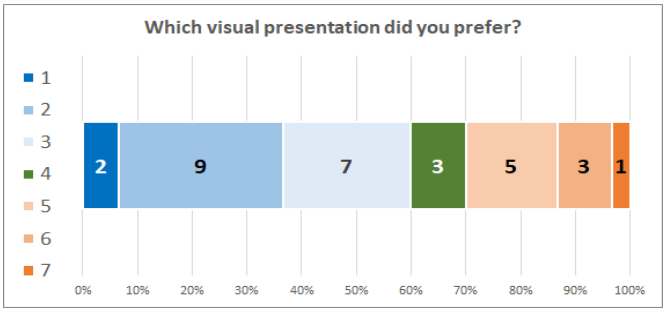

Figure 5: Stacked bar for the ratings giving by users regarding the presentation they preferred.

4 Results and Discussion

4.1 Results

The study consisted of thirty participants; twenty-six male and four female, aged from 20 to 49 years (average 23.3 years old, Stdev = 4.93). Twenty users frequently used video-on-demand and movie-browsing services, eight had some experience and two had no experience at all. Five subjects had frequently used these services on a mobile device, while sixteen only had some experience and nine never used these services on mobile devices at all. The participants were split equally among both the Nexus 9 and Nexus 7 devices, resulting in fifteen subjects for each device.

Usability and Interface (SQ1)

From the ratings given for the comparative questions, we see that subjects strongly prefer the *Standard* interface. For the first question A1, four users preferred the *MovieWall* (rating 5 or higher) and twenty-five preferred the *Standard* interface (rating 3 or lower), with a mean score of 2.63. For the question; "Did the interface and user experience have an impact on your choice?", all participants answered "yes". Looking at the ratings for A2, nine subjects favored the *MovieWall*, while eighteen preferred the *Standard* interface, with a mean score of 3.43. Yet, for question A3, only four participants rated in favor of the *MovieWall*, while twenty-four valued the *Standard* interface more, with a mean score of 2.63. Ratings for A1 and A2 can be found in Figure 4, and ratings for A3 can be found in Figure 5.

The SUS questionnaire show an even stronger preference for the *Standard* interface. For the *MovieWall*, the mean SUS-total settled at 57 out of 100, while the *Standard* interface scored 83 out of 100, that is 45.6% higher than the *MovieWall*. The mean ratings for each SUS statement S1-S10 can be found in Figure 6. Some key points that stood when looking at these statements:

- Statement S8, *"I found the system very cumbersome to use"* was rated very high for the *MovieWall*. Nineteen users agreed with this statement, rating 5 or higher, whereas only seven disagreed, rating 3 or lower. By contrast, only one user rated the *Standard* interface as cumbersome. This participant gave both interfaces the same score (5). The mean scores for this statement are 1.9 and 4.53 for the *Standard* interface and the *MovieWall* respectively. The ratings given are highlighted in Figure 7.
- For all positive SUS statements (statements 1, 3, 5, 7 and 9), the minimum combined score of the *Standard* interface (19 out of 30) was higher than the average of the *MovieWall* (16 out of 30).
- On average, the *Standard* interface scored better on each individual statement (higher for positive statements, lower for negative statements).
- Twenty-eight users gave the *Standard* interface a better SUS-score than the *MovieWall* interface.

Comments from the open interview were utilized to explain the aforementioned results. To analyze these qualitative statements, we accumulated and tallied key words and phrases. Positive characteristics of the *MovieWall* that stood out included:

- Eleven users mentioned that the *MovieWall* was a fun interface.
- Eleven subjects said they found the *MovieWall* a cool concept.
- Ten users said they appreciated the focus on exploration and discovery, though they did not specifically mention they would use the interface for this purpose.
- Four subjects thought using the *MovieWall* was exciting.

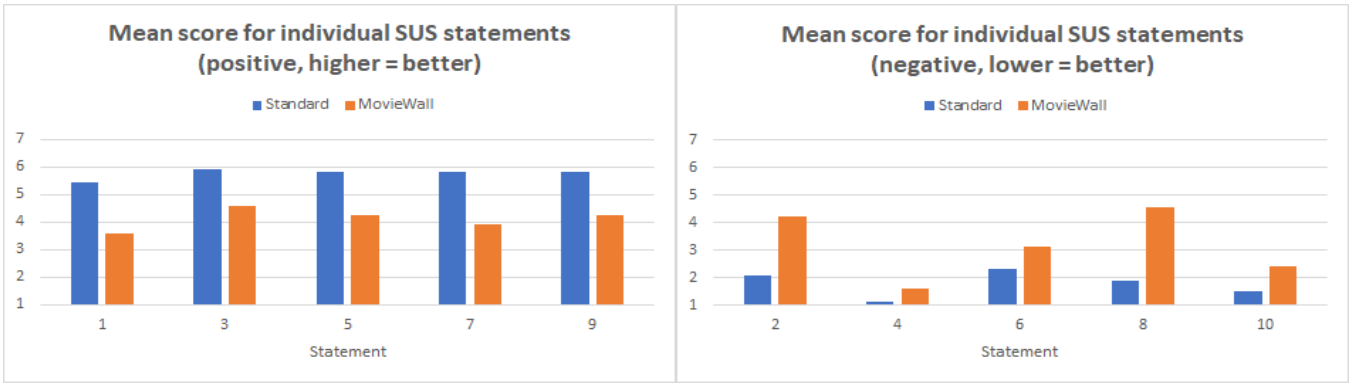

Figure 6: Mean scores for each SUS statement. Positive statements on the left, negative statements on the right.

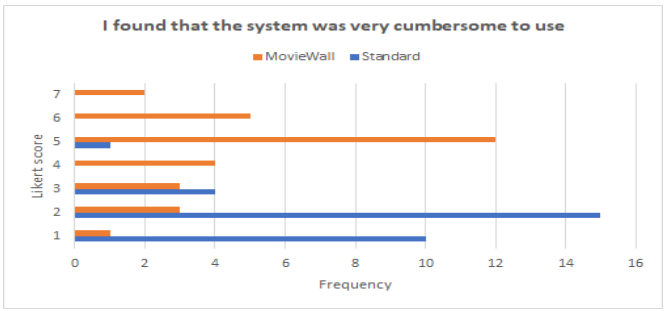

Figure 7: Rating frequencies for the statement *"I found the system very cumbersome to use."*, for both interfaces.

- Four participants specifically mentioned the *MovieWall* had or could have a niche use, hinting at the aforementioned discovery aspect.

Negative characteristics that stood out included:

- Sixteen users mentioned they got lost easily using the *MovieWall*, and specifically noted this as being detrimental to their experience.
- Eleven subjects were negatively impacted by the lack of personalization options. They expected suggestions based on their input.
- Nine participants were either annoyed by the lack of connections between movies and clusters, or confused as to whether there was a connection at all.
- Nine subjects found the current structure detrimental and lacking.
- Seven users rated the current implementation of the *MovieWall* as poor and inadequate.
- Six participants found the *MovieWall* distracting.
- Six subjects had trouble finding movies.
- Five users found the *MovieWall* overwhelming.

Comparatively, no negative characteristics stood out for the *Standard* interface. Frequently mentioned, positive characteristics include the following:

- Fourteen participants stated that the interface felt familiar, with some specifically mentioning specific services (e.g. Netflix).
- Ten users specifically said the *Standard* interface was better.
- Six subjects mentioned that it was more practical.
- Four users stated that the *Standard* interface was more clear.

Overall, statements were favorable towards the *Standard* interface, though some users recognized the potential that the *MovieWall* concept has.

Looking at the results of the logs, eleven subjects did not use the "See More" function and eleven participants did not use the "See All" function of the *Standard* interface. Ten users also did not tap the underlined words in the details panel while using the *Standard* interface at all. Conversely, only four users did not use the "More Like

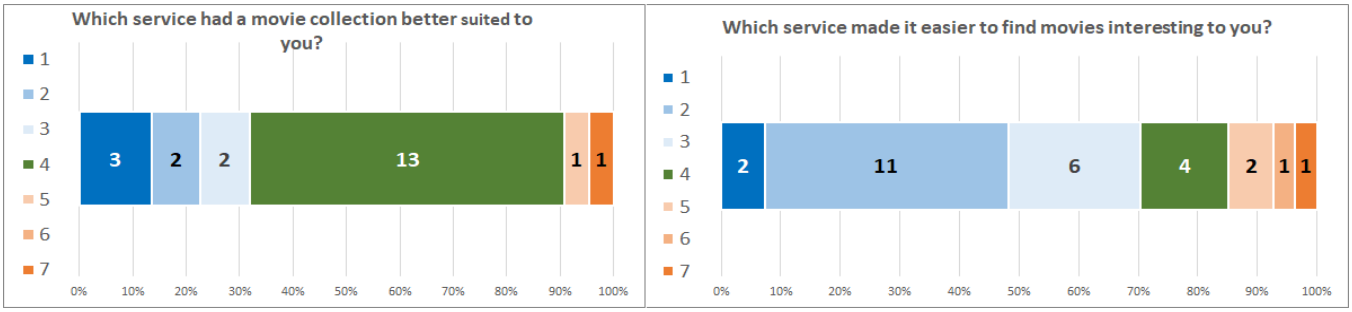

Figure 8: Stacked bar graphs showing ratings given for A4 and A5. On the left, users overwhelming picked the neutral option (4, in green). The figure on right shows more favoritism towards the *Standard* interface regarding ease of finding interesting movies, as shown by the blue portions.

This" feature. The "Home" function remained unused by ten subjects, while the "Back" button was used by everyone, bar two users. By contrast, nearly all functionality was used by all users in the *MovieWall* with one exception: the filter functionality, activated by tapping on the underlined words in the details panel, remained unused by five participants.

Regarding frequently requested features observed during the experiment and the interview, users often attempted to tap the descriptive headers as seen when fully zoomed out in the *MovieWall* interface, and were met with disappointment once they realized such a feature did not exist. Many users also pointed out that they expected more guidance in the *MovieWall*, so they did not have to zoom out to regain their whereabouts. Other missing functionality included the lack of a search function on both interfaces.

Database and Collection (SQ2)

For the question; "Did the movies that you saw have an impact on your preference?", only seven participants answered "yes". Question A4 was rated by twenty-two of the thirty users. The mean score for this question was a 3.5, with two favoring the *MovieWall* and seven preferring the *Standard* interface, and thirteen remained neutral. Conversely, comparative question A5 was answered by twenty-seven users, with nineteen favoring the *Standard* interface and four preferring the *MovieWall*. The mean rating for this question was a 3.0. Ratings for questions A4 and A5 can be found in Figure 8.

Screen Size (SQ3)

As scores for the SUS questionnaire and each comparative question A1-A5 appeared to be slightly different among the Nexus 9 and Nexus 7 groups, we performed two-sample, two-tailed t-tests between the two to check the significance. For each tested score, we expect no significant difference between the two groups. As such, we confirm that there is no difference should the null-hypothesis, "There is a difference in scores", be omitted. Results of all tests can be found in Table 2. For both the *MovieWall* and *Standard* SUS-score comparisons, the null-hypothesis was omitted. The null-hypothesis was also omitted for every comparative questions, except for A5. Thirteen users of the Nexus 9 group and fourteen subjects of the Nexus 7 gave a rating for this question. For A5, the table shows a large variance between the two groups (2.47 and 0.64) and the mean scores given by the Nexus 9 and Nexus 7 groups are 3.85 and 2.21 respectively. For all other ratings, mean scores were much closer together, as seen in Table 2.

Other observations of interest were also tested. However, they were inconclusive. These include:

- The Nexus 9 group gave a mean score of 4.8 for SUS statement S5 (*MovieWall*), while the Nexus 7 group gave this statement a mean score of 3.73.
- The mean scores for SUS statement S6 (*Standard*) were 2.73 and 1.87, as rated by the Nexus 9 group and Nexus 7 group respectively.
- For SUS statement S10 (*Standard*), the Nexus 9 group rated a mean score of 1.93. Comparatively,

T-test	Equal Variance?	μ_1	μ_2	σ_1^2	σ_2^2	d.f.	t_{stat}	t_{crit}
SUS (*Standard*)	Yes	81.22	84.89	60.90	43.44	28	-1.39	2.05
SUS (*MovieWall*)	Yes	85.33	87.00	732.02	405.54	28	-0.19	2.05
A1	Yes	2.53	2.73	2.41	2.21	28	-0.36	2.05
A2	Yes	3.53	3.33	2.41	3.10	28	0.33	2.05
A3	Yes	3.00	3.25	2.14	2.02	25	-0.45	2.06
A4	Yes	3.80	3.25	1.96	2.02	20	0.91	2.09
A5	No	3.85	2.21	2.47	0.64	18	3.36	2.10

Table 2: Results from two-tailed t-tests. All tests were performed using $\alpha = 0.05$.. All data was tested using an F-test to check for equal variance.

participants of the Nexus 7 group rated an average of 1.07.

4.2 Discussion

Going back to SQ1; "How do the two interfaces compare in terms of usability for a general browsing task with vague search goals?", it is clear that the *MovieWall* implementation appears to be severely lacking after looking at the results of the SUS questionnaire and comparative questions A1-A3. In particular, nineteen users found the *MovieWall* cumbersome, and the SUS score for the *Standard* interface was 45.6% higher. Furthermore, twenty-five users preferred the *Standard* interface over the *MovieWall*, as seen from the ratings given for A1. This is further emphasized by the answers given for A3, where twenty-four favored the *Standard* interface.

Several reasons exist for these results. Sixteen users stated getting lost easily and eleven users disliked the lack of personalization options. This is further emphasized by the log results; the "See All" and to lesser degree the "See More" features were introduced to provide users an easy way to find every movie in the database without using a query. Yet, eleven users did not use the "See All" feature, and eleven did not use the "See More" feature. Even though these features were underused, the ratings were still overwhelmingly in favor of the *Standard* interface.

Unique, positive characteristics of the *MovieWall* identified by the open interview include a fun browsing experience, as noted by eleven users, and the emphasis on discovery, as noted by ten users. These are defining strengths of the *MovieWall* which the *Standard* interface does not possess.

Looking at the results of the question; "Did the movies that you saw have an impact on your preference?", and the ratings given for A4 and A5, we answer SQ2; "How does the interface influence the perceived quality and size of the explored database?", as follows: While users did not see a difference in the database, the *Standard* interface gave them a better time finding movies of interest. Twenty-three participants stated they were not influenced by the movies they had seen when selecting their preferred interface. Additionally, the results of A4 indicate that most users did not see the movie collection of each interface as being different, though individual preferences exist; for this question, seven users preferred the *Standard* interface and two picked in favor of the *MovieWall*. Ratings given for A5 show that the *Standard* interface made it easier to find movies of interest; nineteen users showed a preference for it, while only four subjects favored the *MovieWall*. Given the results of SQ1 and A5, we argue that the small favor for the *Standard* interface in the ratings given for A4 may be due to better usability. This further indicates that participants did not have a different perception of the database itself.

Comparing results for the Nexus 7 and Nexus 9 groups show no significant differences, with the exception of A5. However, only twenty-seven ratings were given for this question in total, and variances differed greatly ($\sigma_1^2 = 2.47$, $\sigma_2^2 = 0.64$) As such, the question SQ3; "Does screen size have a significant influence on any of the aforementioned aspects?" with "No". Several findings, such as the differences for S5 (*MovieWall*), S6 (*Standard*), and S10 (*Standard*), suggest that both interfaces may perform better on larger devices than smaller ones. However, these findings are inconclusive, as no significant differences were found.

5 Conclusion and Future Work

Looking back at our main question; "Is the concept of the *MovieWall* a complement to existing state-of-the-art approaches and could it even replace them completely?", results indicate the answer to be "no", given the current implementation of the *MovieWall*. Despite some positive aspects and the outcome of Nefkens' pilot study [8], our comparative experiment showed that for the given task, users have a clear preference for the *Standard* interface implementation. Yet, although minor, positive aspects about the *MovieWall* have been observed as well, and there might be evidence suggesting that at least some of the negative results could be contributed to the current implementation or the study design.

From the SUS scores, it is clear improvements to usability and user experience are of utmost priority. In the interview, many subjects stated they got lost easily while using the *MovieWall* interface, which negatively impacted their experience. New implementations should therefore grant the user some form of guidance, so they know where they are within the *MovieWall*, without resorting to zooming out.

Another point that stood out was the demand for customization and personalization, as seen from the interview statements. Users expected their input to influence the *MovieWall* interface by changing the arrangement of movies. Future implementations that allow users to remove movie genres which are not of interest to them may solve this problem, improving user experience. Allowing users to create new clusters through some form of filtering may have the same effect, as it gives them more control over the content presented.

One key problem regarding user experience of the *MovieWall* implementation was its structure. Results from the interview indicate many users are negatively impacted by the current lay-out or having trouble finding a connection between the movies they see. This is largely caused by movies changing discretely from one cluster to the other; moving from the "Action" cluster to the "Adventure" cluster does not guarantee seeing movies that are both "Action" and "Adventure".

The current user study was short-term and forced users to pick between the two interfaces. The former may have pushed participants into picking that which is more familiar to them; two users specifically stated that the *MovieWall* "grew onto them" in the interview. The latter emphasized on competition between the two, rather than looking at the *MovieWall* as a complementary function. While results indicate the current *MovieWall* implementation to be inadequate, future research may focus more on implementing both interfaces together, allowing users to switch between the two. Future studies would then emphasize less on making a forced decision between the two, highlighting the *MovieWall* concept as a complementary function.

While we believe state-of-the-art movie browsing applications may not be the best option for exploring movie collections, the current *MovieWall* implementation does not provide a better experience for this task. Though the concept still shows promise, the aforementioned issues must be answered before it can be verified as a suitable alternative.

References

[1] Bobadilla, J., Ortega, F., Hernando, A., and Bernal", J. "a collaborative filtering approach to mitigate the new user cold start problem". *"Knowledge-Based Systems" "26"* ("2012"), "225–238".

[2] Brooke, J., et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry 189*, 194 (1996), 4–7.

[3] Carvalhal, S., et al. Active exploration of online movie databases by visual navigation. Master's thesis, 2017.

[4] Glowacka, D., Ruotsalo, T., Konuyshkova, K., Kaski, S., Jacucci, G., et al. Directing exploratory search: Reinforcement learning from user interactions with keywords. In *Proceedings of the 2013 international conference on Intelligent user interfaces* (2013), ACM, pp. 117–128.

[5] Kangasrääsiö, A., Glowacka, D., and Kaski, S. Improving controllability and predictability of interactive recommendation interfaces for exploratory search. In *Proceedings of the 20th international conference on intelligent user interfaces* (2015), ACM, pp. 247–251.

[6] Marchionini, G. Exploratory search: from finding to understanding. *Communications of the ACM 49*, 4 (2006), 41–46.

[7] McSherry, F., and Mironov, I. Differentially private recommender systems: Building privacy into the netflix prize contenders. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining* (2009), ACM, pp. 627–636.

[8] Nefkens, M. Moviewall: A novel interface concept for visual exploration of large movie collections. Master's thesis, 2017.

[9] Pang, P. C.-I., Verspoor, K., Chang, S., and Pearce, J. Conceptualising health information seeking behaviours and exploratory search: result of a qualitative study. *Health and Technology 5*, 1 (2015), 45–55.

[10] Wei, J., He, J., Chen, K., Zhou, Y., and Tang, Z. Collaborative filtering and deep learning based recommendation system for cold start items. *Expert Systems with Applications 69* (2017), 29–39.

[11] Wildemuth, B. M., and Freund, L. Assigning search tasks designed to elicit exploratory search behaviors. In *Proceedings of the Symposium on Human-Computer Interaction and Information Retrieval* (2012), ACM, p. 4.

Annotated Appendix

3.1 Appendix A: Literature review

The following shows the literature review done at the start of this thesis project. The results of this work formed the basis for the research leading to the scientific paper.

Introduction

With the introduction of the Web and an unmeasurable amount of information publicly available, methods of searching information were quickly implemented. Indeed, most users of the Internet regularly utilize search engines in some form, retrieving information through simple or complex lookups. With the immense amount of information growing and collections becoming ever larger, content providers collectively moved towards recommender systems, which select and provide content for their users without the need of lookups. However, recommender systems have their own set of problems, including the infamous 'cold start' problem. Solutions to the cold start problem often provide their own challenges. Meanwhile, many researchers have moved their attention back to exploratory search, a method often used unconsciously and widely adopted before the introduction of the Web. Believing exploratory search may provide solutions to the challenges introduced by look-ups and recommender systems, recent research has made various implementations while isolating important factors of exploratory search. In this literature review, we look at the dismissal and resurgence of exploratory search in relation to search engines (look-ups) and recommender systems, the challenges and recent developments of each. Finally, we identify points of interest for future research, formulating these in the form of research goals.

Look-up and Search Engines

Searching using search engines has been widely adopted for a long time. Most users of the Internet know a few of these, like *Google*, *Bing* and *Yahoo!*. Many

providers of a large collection of goods also incorporate a (form of) a search engine, such as YouTube. Search engines perform extremely well when the exact name of the desired item is known as it is looked up, while advanced engines can even return the item as long as the query relates well enough to the item in question. Still fairly new and ever evolving, search engines possess many challenges and developments. Henzinger et al. describe such challenges as filtering spam entries, managing content quality, evaluating quality, creating and maintaining web conventions, removing duplicates and handling vague-structured data [7]. Cambazoglu defines further challenges regarding the scalability of search engines[4]. Chang et al. specifically define advances and challenges for semantic image/video search [6]. Other research includes the use of different types of inputs, such as used in music search engines[14].

One obvious trait of search engines is the input required to be of any use. Naturally, a query requires a formulation to be of any use. Additionally, such a query needs to be interpretable by the search engine. Finally, there can always be semantic differences between the user and the system. Accommodating to a single user is relatively simple. However, adapting to a much larger audience, especially one that is culturally diverse, can prove problematic. One such cultural example includes the use of the bag of words query, semantics and natural language query: does looking for an 'orange monkey' return a monkey that is orange, a monkey and an orange, or an orangutan?

Recommender Systems

To combat the shortcomings of look-ups and engines, recommender systems adapt to individuals who do not possess the expertise to formulate proper queries that would generate their desired outcome [19]. Recommender systems instead present items of interest to each user, consisting of both personalized choices and general items (e.g. well-performing articles). Like look-ups, recommender systems score items depending on a set of criteria, presenting the user with items that score highest on this set. Information relating to users may be obtained explicitly or implicitly. An example of explicitly obtained information includes asking a user whether they liked the article they purchased recently. Information obtained implicitly is gathered by interpreting the user's actions, such as navigating certain products. Using this information, a recommender system presents users with items that are related to the user's behavior, narrowing down the large number of items to a much smaller collection actively present, coping with information overload. Even before the web, recommender systems already existed informally and were used at large. Examples include racks of items in a record shop containing the top 10 best selling songs, or the shop clerk recommending products based on the contents of your shopping cart.

Formally, Ricci et al. define the goals of a recommender system from a service provider's perspective as follows: increase the number of items sold, sell more diverse items, increase user satisfaction, increase user fidelity, better understanding of what the user wants [19]. These are collectively defined as increasing the conversion rate. Ricci et al. further define the goals from a

user's perspective, which consist of: finding some good items, finding all good items, annotation in context, recommend a sequence, recommend a bundle, just browsing, finding a credible recommender, improving the profile, self-expression, helping others and influencing others.

Recommender systems utilize various different techniques to generate recommendations [19]. Demographic profiles separate users on known traits and present content appropriate to each demographic, while knowledge-based systems base their recommendations on the needs of the user and the features of the items. Content-based recommenders match items in terms of similarity and present these to the user [17], while collaborative filtering systems utilize past events from users with similar tastes to present recommendations to the current user [20]. Many systems utilize hybrid forms of these. From the above, it is clear that recommender systems still require input initially to generate a profile that the system can utilize [19]. This problem is widely defined as the 'cold start' problem. Initially, explicitly obtained information can alleviate this issue, though this may be intrusive to the user and consist of very long questionnaires to obtain proper information depending on the rating system and content in question. Additionally, interaction from the user is also a large challenge, ever researching attributes of user behavior, their importance and the relationship to their choices, as well as efficient and pleasant feedback towards the system. More challenges appear when attempting to recommend other things than just related items, while privacy when observing user behavior is constantly put into question.

The cold start problem may be the most prominent problem, with many researchers attempting to solve or alleviate it. The works of Lam et al. show a hybrid model using collaborative and content data to present recommendations to novel users, utilizing the idea of users with similar features enjoying the same content [10]. Lam et al. specifically use age, gender and job to predict recommendations, using a Bayesian Network to learn user preferences among the training set, then matching an item to each user in the test set. Note that the work still uses a form of explicitly provided information to address the issue, with information that may be regarded as sensitive (job). Ahn addresses the cold start issue through the use of a similarity function, PIP, specifically for collaborative filtering recommenders [1]. Pereira et al. also use learning along with co-clustering to address the problem, splitting the cold start problem between 'pure' (zero ratings) and 'incremental' [18]. Their method utilizes matrix factorization K-nearest neighbors, using side information to estimate features of a user with a learned model and subsequently mapping these to existing users from training, which functions very similar to the works of Lam et al.

It is clear that recommender systems provide many benefits over simple or complex lookup. As such, even large search engines such as Google often adapt search queries to the user depending on the information available, creating a hybrid between a look-up and a recommender system, to present more appropriate content. Recommender systems will in many ways continue to suffer from the cold start problem, as well as knowledge intrusion, whereas lookups still have their challenges in user proficiency. However, both have clear advantages and

use-cases compared to one another, with both systems often coexisting in interfaces in the form of complementing features.

Exploratory Search

Another system that was informally prominent before the creation of the Internet exists in the form of exploratory search. White and Roth formally define exploratory search as follows: "Exploratory search describes an information-seeking problem context that is open-ended, persistent, and multifaceted, and information-seeking processes that are opportunistic, iterative, and multi-tactical. Exploratory searchers aim to solve complex problems and develop enhanced mental capacities. Exploratory search systems support this through symbiotic human-machine relationships that provide guidance in exploring unfamiliar information landscapes." [22] Mirizzi and Noia further highlight three important traits of searching that exploratory search attempts to answer: "learning, discovering and understanding novel knowledge on complex and sometimes unknown topics."[13] Many researchers acknowledge that scenarios in which a user may require this kind of information retrieval may not be answered by look-ups and recommender systems. Pang et al. illustrate on such example in their work: if patients or their cares don't know the right keywords for their health problems, how can they effectively find useful health information [16]? Indeed, before users had search engines and a wealth of information directly at their disposal, users would often find information and knowledge by pure chance or thoughtful exploration. An example of pure chance portrays a customer walking through a farmer's market, finding and purchasing a type of fruit they never imagined existed. Examples of thoughtful exploration include a video game fanatic browsing the local video game store's entire collection, obtaining knowledge from the video games that were made and more. Even window shopping can be considered a form of exploratory search, where the goal may be the browsing of items itself rather than the selection of items.

In terms of pleasure and satisfaction, these examples can be related to the 'Discovery' aesthetic as defined in the widely accepted MDA framework for game design [8]. In fact, reformulating the goals of exploratory search, one could state that discovery and exploration of elements is inherently part of the fun of exploratory search. Marchionni further argues that "we seek to fulfill social and psychological needs to belong and to know our world" as being an important reason for exploratory search to exist [12].Marchionni also highlights the need for interaction in search. Dynamically updating queries and displays, judging returned content and revising the query, highly interactive user interfaces are at the forefront of turning people into "pioneers and adventurers in a new world of information riches awaiting discovery along with new pitfalls and costs." [12] He defines the goals of Exploratory Search in two activities, 'Learn' and 'Investigate', which are divided further into many other subcategories. Note that these goals are largely akin to the traits illustrated by Mirizzi and Noia, with 'investigate' encompassing 'discovering' and understanding'. Some earlier research by Kules et al. also suggests that exploratory searchers mostly look at

the results of a faceted search, spending less time on the facets and almost no time on the query[9]. Specifically, users would be fixated on the results.

By looking at the differences between exploratory search and traditional look-up, tasks can be defined in which an exploratory search system should excel and where look-up should be better. Defining these tasks can also help evaluate an exploratory search system in a quantitative manner. Wildemuth and Freund provide definitions specifically for exploratory search tasks, guiding researchers during investigation [23]. In their work, they define exploratory search tasks to have the following characteristics:

- Exploratory search tasks are open-ended. There is no specific answer.
- Exploratory tasks are associated with the goals of learning and investigation.
- Tasks are general rather than specific. There is a very abstract goal.
- Tasks target multiple items. There are many possible solutions to the task.
- There is a degree of uncertainty.
- The tasks are multi-faceted, including multiple aspects or concepts, incorporating multiple subtasks.
- The process is dynamic, as the user's understanding of the topic gradually changes or motivations behind the task shift.
- The process occurs over time, having a fairly long session.
- The task is accompanied by other information and cognitive behaviors, e.g. sense making and decision making.
- The problems that elicit exploratory search tend to be ill-structured.
- The tasks are fairly complex, "not too easy".

Athukorala et al. use the definition of search categories as defined by Marchionni, as well as the characteristics of exploratory search tasks as defined by Wildemuth and Freund, to better define different subcategories of exploration and gain understanding of how users behave in exploratory search[2]. In their work, lookups tasks are defined as finding facts to a specific questions (query), which are characterized by precise search goals with simple search paths. Lookup tasks involving thinking or understanding are referred as interpretive tasks, being more focused and goal-oriented than exploratory tasks. In contrast, their work defines exploratory tasks as open-ended and imprecise, without a single answer accomplishing the user's information needs and no clear criterion for the end of the search. Note how this definition largely correlates to the definition made by Wildemuth and Freund. The authors define six tasks by name: knowledge acquisition, planning, comparison, fact-finding, navigation

and question answering. The first three tasks are defined as exploratory and the last three tasks are defined as look-up tasks, with the third task in each set being 'borderline', largely differing from standard tasks in each set. With exploratory search being researched extensively, novel implementations are becoming more prominent, answering the needs and goals of users as described before. The visual berry-picking method defined by Low et al. utilizes an interactive approach with a highly faceted and large information space [11]. Their work creates the impression of panning a large global map through item neighborhoods, where the currently selected item generates clusters with related items nearby. General responses are positive, with users able to handle the new interface, though the authors note bias towards the interface that was presented first. A different implementation by Thudt et al. concentrates around serendipity and spontaneous discoveries, which aligns well with multiple aspects of exploratory search [21]. Their implementation, the Bohemian Bookshelf, utilizes playful approaches, visual connections and abstraction, while offering flexibility and accessibility between elements, to stimulate and encourage serendipity. The Bohemian Bookshelf too was a large success among subjects, idealizing the thought of "maximizing the number of possibly relevant objects" over "minimizing the number of possibly irrelevant objects". The authors do note that distraction through complexity, which may be beneficial to exploratory search. However, the authors also note that users requested features that emphasized on targeted search strategies over exploratory behavior. Further work by Carvalhal emphasizes on user engagement and active exploration through visual navigation [5]. Akin to the approach by Low et al., users move from item to item by selecting interesting metadata presented by each item, effectively navigating the system solely through visual representations. Once again, users did not perceive the interface being any more difficult to more traditional implementations, while being more intrinsically motivated. Finally, the implementation by Nefkens, experiments with visual exploration through the use of the *MovieWall* [15]. The *MovieWall* consists of a large grid with many movie posters clustered together, with users panning a mobile tablet device to navigate the interface and selecting items of interest to gain more information on. Results show that users were generally positive and enjoyed the interface, though the phenomenon of being overwhelmed was a real issue. Other criticism includes the dependence on movie posters for further decision making. Other experiments using the *MovieWall* concerning randomization of items within clusters showed that aggressive randomization was not appreciated, with users favoring standard clustering or light randomization.

Conclusion

Exploratory search appears to have various reasons to be implemented and encouraged, while having clear weaknesses where a different system may be more suitable or requiring a hybrid implementation to shore up said deficiencies. Most notably, implementations that show clear relationships between elements and encourage discovery and understanding appear to be accepted by test subjects.

The relationship between look-up and exploratory search are not fully clear yet, with the strengths and overlaps of each system compared to the other still being researched. However, when comparing recommender systems to exploratory search systems, much less is known. This is even more prominent when including systems that utilize hybrids or combinations of recommender and look-up systems, such as YouTube. Where does each system excel, and where do they overlap? What systems do users prefer in which situation? Can one be a substitute to the other? Can they complement one another? Answering these questions should provide insight into the use-cases and potential developments of exploratory search, especially in relation to recommender systems and search engines, as well as the strengths and weaknesses when directly competing with each other.

3.2 Appendix B: Study Implementation

This section contains an overview of the application used for the comparative user study. The application contains both the *MovieWall* interface and the *Standard* interface. The first subsection contains the initialization, which is important for the *Standard* interface and may also be useful for future implementations of the *MovieWall*. The next subsection shows the *Standard* interface in detail. The last subsection presents the differences, attempted features and future challenges for the *MovieWall*.

Initialization

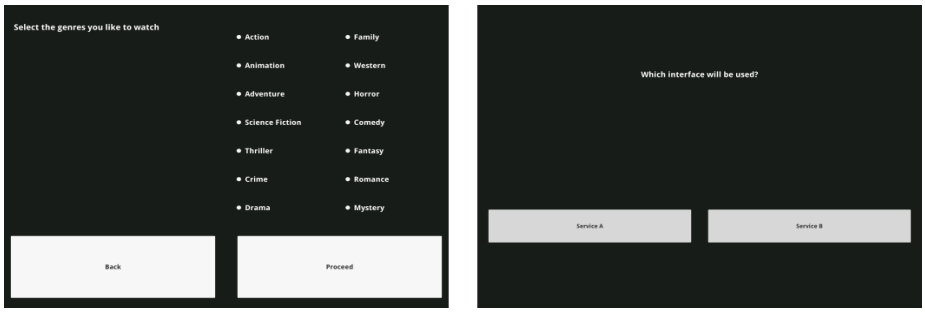

Figure 3.1: On the left, genre selection. On the right, service selection.

To simulate a recommender system based on a test subject's preference, usage of the application starts by selecting the genres that a user prefers, from the set of genres listed in Figure 1 (left side). These genres are stored and used in the *Standard* interface. Afterwards, users can select the interface they wish to use in the interface selection screen (Figure 1, right side), where "Service A" will lead them to the *Standard* interface and "Service B" leads them to the *MovieWall*. Hitting the Android back button will show a pop-up that allows the user to quit the current interface, returning to the interface selection screen. .

Standard Interface

The *Standard* interface is based on the state-of-the-art interfaces known from online movie services such as Netflix. These applications typically create multiple horizontal lists, which each containing items based on a filter that is described in the header. For example, a horizontal list containing action movies will typically have a header saying "Action" or "Movies with the Action genre". The items in the list as well as the lists themselves are then sorted based on some value, such as popularity. These horizontal lists are the main aspect of the *Standard* interface. In state-of-the-art interfaces, filters used can become increasingly complex based on observed patterns in user behavior. For our implementation, we only

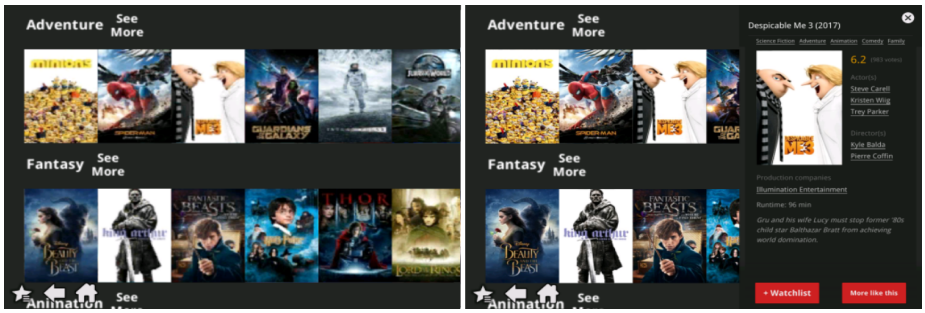

Figure 3.2: The home screen of the *Standard* interface is shown on the left. On the right, the same view with the details panel opened.

utilize information about the movies and the preferred genres of the user, which minimizes the amount of private data we store for each user. As the *MovieWall* does not utilize user data, this allows for a fairer comparison between the two interfaces, as well as challenging each interface to perform adequately without the use of elaborate algorithms and extensive knowledge of users.

For all horizontal lists, the items within each list are produced as follows. First the filters used are applied. If no entries return, the list will be removed. If there are entries, they are sorted from most to least popular, and the first thirty entries are inserted into the list. In some views, if a previous list already contains an entry, duplicates are removed to provide a better experience (and if no entries remain, the list is removed as before). On the interface, each list shows up to six entries at first, and can be dragged horizontally up to thirty entries. Dragging vertically anywhere on the screen scrolls through all the lists, with up to three lists being in view at any point in time. Users start at the home screen, which can be seen on the left of Figure 2. On the home screen, all genres listed in Figure 1 are used as filters for the lists. The lists themselves are sorted so that the genres preferred by the user appear first, followed by the remaining genres. This provides a naive method of presenting recommendations, which avoids showing random entries at the start. To return to the home screen, users can press the home button, the little house icon in the lower-left corner, which is accessible at all times. From the home screen, users can tap on "See More" near the headers to go to a view similar to Figure 3, using the genre next to "See More" as a reference. Movie posters can be tapped to bring up a movie details panel identical to the one in the *MovieWall* with some functionality differences as seen in Figure 2, on the right side. First, users can tap the "More Like This" button to produce the view presented at Figure 4. Pressing any of the underlined words instead produces a drastically different view as seen in Figure 5.

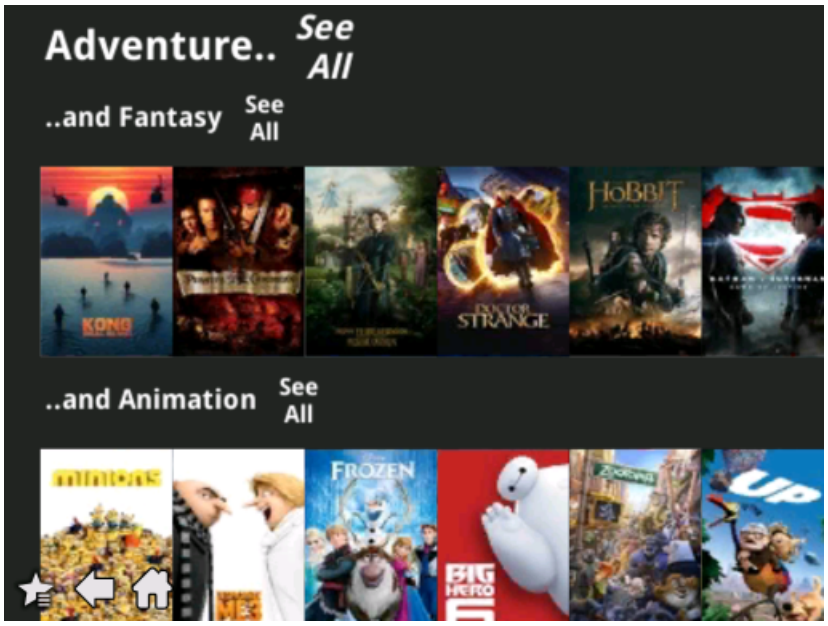

Figure 3.3: Illustration of the "See More" function.

In the view presented by Figure 3, the "See More" text near each header in Figure 2 is replaced by "See All". Tapping on "See All" produces the view shown in Figure 5. A header is placed at the top as well, showing the genre used as reference to the user. The horizontal lists created each use the genre described at the top, as well as another genre from the initial selection of genres, resulting in up to thirteen lists depending on whether entries with both genres exist. Tapping on "See All" next to each list header will use both filters as reference for the view in Figure 5, while tapping on "See All' at the top of the screen will only use that genre. This view allows the user easy access to combinations of genres they might find interesting, which is a common feature in state-of-the-art interfaces.

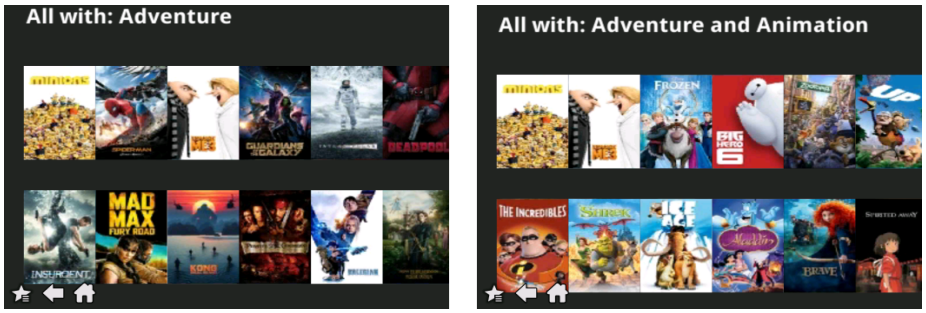

Figure 3.5: Illustration of the "See All" function. On the right, a single filter is applied. On the left, use of two filters is showcased.

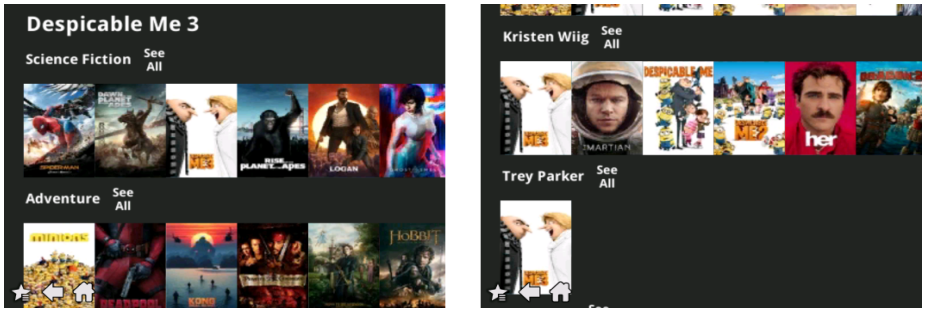

Figure 3.4: Illustration of the "More Like This" function. On the right, the view is vertically dragged to show actor filters have been applied.

The "More Like This" function allows users to find movies similar to the ones that interest them more easily akin to state-of-the-art interfaces. The view in Figure 4 is similar to Figure 3, with a few key differences. First of all, the top header does not contain a "See All" button, as the method of combining would be confusing and illogical to the user. Second, the reference used is a movie instead of a genre, as the user would naturally assume when tapping "More Like This". As a result, each horizontal list produced uses either a genre, an actor, a director or a company as filter. The genres, actors, directors and companies used can be seen in the details panel, prior to tapping the button. The lists are sorted so genres appear first, actors second, directors third and companies last. Tapping on "See All" uses that filter as reference for the view in Figure 5.

The view shown in Figure 5 was added to provide users with full access to the collection as conveniently as possible. This makes a comparison with the *MovieWall* more fair, as it provides full access to any genre quickly. This view differs drastically from the other views. It shows a descriptive header at the top, indicating what filter was used. This filter is then applied and items are sorted by popularity, akin to the other views. All entries are then inserted in a grid-like

fashion: up to six movies horizontally and expanding vertically. Users can drag vertically in order to show all these entries; horizontal dragging is disabled and does not have a function here.

Users can undo their actions by pressing the back button, the arrow icon near the home button. This will reload the previous view. All other functionality in this interface is similar to the *MovieWall*. Users can add movies to the watch list which can be viewed at any point by tapping on the star icon.

MovieWall

For the comparative user study, the *MovieWall* was largely kept the same as the iteration tested in [15]. After some initial reactions, one major change was made for the comparative user study; users start with the interface fully zoomed out, showing the headers and location of each cluster. This change was made as initial reactions often mentioned being unaware of zooming (which was also present in [15]) and preferred being able get an overview of the lay-out before continuing, given the short task duration. All entries within each cluster were combined without any randomization, as this provided a better experience during the tested iteration compared to heavy randomization [15].

One issue of the *MovieWall* we investigated was its clustering. Currently, this process takes a significant amount of time when attempting to make neat clusters. Clusters are based on Euclidean distance, which makes them fairly round in form and difficult to fit within a rectangular space. To improve the fitting, spaces are optimized iteratively to fill the entire grid while keeping the entries of each cluster as close to the center as possible. While this does not prove to be an issue offline on a consumer desktop, it causes significant slow-down on a mobile device, which prevents such a feature from being done in real-time. Other issues include the device heating up significantly, as well as consuming a lot of battery power. All these effects would impact user experience negatively. Another problem stems from the resulting clusters, which are still fairly stretched-out even after optimization. This issue may be handled more elegantly if the interface itself is more dynamic: users do not need to see entries of a cluster that is far away. Odds are these entries do not need to be loaded until the user is about to enter a different cluster. However, the benefits of solving these problems are unclear, as results from [15] had shown that the *MovieWall* would be adequate as a feature even when the clusters are static. As such, we did not pursue this any further. Given the results of the comparative user study, we believe that solving this problem is now of utmost priority. Improving the structure would bring the *MovieWall* to acceptable standards, so we recommend future research to investigate this problem.

3.3 Appendix C: Comparative User Study

The comparative study summarized in the scientific paper used the implementation described in detail in Appendix B. In the following, we take a closer look at the used methodology, along with a justification of it, and provide further details about the gathered data.

Methodology

The data was gathered in three ways. First was the online form, which consisted of a consent form, SUS questionnaire and additional comparative statements. Next, users were interviewed to provide an explanation behind their choices, as well as probe them for further information. Finally, all actions performed during use of the applications were logged to check the usage of features.

Users were requested to perform the task described in the scientific paper for four minutes, to avoid fatigue and boredom when performing the task twice. This task duration still presented them with enough to time to both adapt to each interface and fully experience each of them. In practice, this time limit was not enforced; users could stop sooner if they felt they experienced enough of the interface, or were stopped by the observer later if participants appeared to be very immersed. As a result, most users took between two-and-a-half and five-and-a-half minutes, with two notable exceptions: One participant used the *Standard* interface for only 62 seconds and the *MovieWall* for 200 seconds. The other participant used the *Standard* interface for only 100 seconds, and the *MovieWall* for 130 seconds.

Online Form

To obtain quantitative data regarding usability and perception of the database, we made an online form consisting primarily of ratings. A SUS questionnaire formed the basis for comparing usability, as it is a convenient and widely recognized questionnaire [3]. Additional comparative statements were added, which allowed users to rate their preference in interface, user experience and database perception. All statements can be found in Table 3.1.

The online form was made using Google Forms. Users would fill their email address at the start to provide consent, of which a copy would be sent to them digitally after completion of the form together with their answers. Users also provided additional information about their age, sex, usage of (mobile) movie-browsing applications. Users would then proceed and follow the flow listed below:

1. Read the task
2. Perform the task on the *Standard* interface
3. Fill the SUS questionnaire for the *Standard* interface
4. Perform the task on *MovieWall* interface

5. Fill the SUS questionnaire for the *MovieWall* interface
6. Answer additional comparative statements
7. Submit the form, finalizing their consent

The SUS questionnaire used the default template as can be found in [3]. SUS statements were rated on a 7-point Likert-scale instead of a 5-point Likert-scale for the sake of consistency with the additional comparative statement, as these were rated on a 7-point Likert-scale to provide enough options to participants. Scores were later normalized to provide a basis for SUS score ratings.

Interview

After completing the questionnaire, users were asked general questions regarding their experience, leading to discussion and statements for each interface. These statements were noted down into raw format and later tallied, grouped together if statements were similar enough. While tallying, references to the statements tallied were noted on the side as validation. The interview provided complementary, qualitative information which helped explain the results of the online form, as well as presenting opportunities for improvement.

Logging

All actions which triggered features were logged into the system. These can be seen in Figure 2. Logged actions were timestamped and, depending on the function, additional data was logged with it. This data was used to look at which features subjects utilized often and which remained unused, and look at any correlations between feature usage and the scores given. For the study, we only looked at unused features, as these proved to be the most valuable.

Data

The figures below show all data gathered, in summarized format.

Table 3.1: All statements and questions used in the online form. S1-S10 are SUS questionnaire statements, A1-A5 are additional statements, Q1 and Q2 are general questions. All SUS questionnaire statements and additional statements are rated on a 7-point Likert-scale.

	Statement / Question
S1	I think that I would like to use this system frequently.
S2	I found the system unnecessarily complex.
S3	I thought the system was easy to use.
S4	I think that I would need the support of a technical person to be able to use this system.
S5	I found the various functions in this system were well integrated.
S6	I thought there was too much inconsistency in this system.
S7	I would imagine that most people would learn to use this system very quickly.
S8	I found the system very cumbersome to use.
S9	I felt very confident using the system.
S10	I needed to learn a lot of things before I could get going with this system.
A1	Which service did you prefer overall?
Q1	Did the interface and user experience have an impact on your choice?
A2	Which visual presentation did you prefer?
A3	Which service gave you a better experience?
Q2	Did the movies that you saw have an impact on your preference?
A4	Please rate which service had a movie collection better suited to you.
A5	Please rate which service made it easier to find movies interesting to you.

Table 3.2: Personal data and device used for each participant.

P	Age	Sex	Former experience?	With mobile device?	Device used
1	22	M	Yes, but few	Yes, but few	Nexus 9
2	49	M	No	No	Nexus 9
3	24	M	Yes, freq.	No	Nexus 9
4	21	F	Yes, freq.	Yes, freq.	Nexus 9
5	23	M	Yes, freq.	Yes, but few	Nexus 9
6	21	M	Yes, freq.	Yes, but few	Nexus 9
7	21	M	Yes, but few	Yes, but few	Nexus 9
8	24	M	Yes, but few	Yes, but few	Nexus 7
9	25	M	Yes, freq.	Yes, freq.	Nexus 7
10	22	M		Yes, but few	Nexus 7
11	22	F	Yes, freq.	No	Nexus 7
12	20	M	Yes, but few	No	Nexus 7
13	22	M	Yes, freq.	Yes, but few	Nexus 9
14	23	M	Yes, freq.	Yes, but few	Nexus 7
15	21	M	Yes, freq.	Yes, freq.	Nexus 7
16	20	M	Yes, freq.	Yes, but few	Nexus 7
17	23	M	Yes, freq.	Yes, but few	Nexus 7
18	23	M	Yes, freq.	No	Nexus 7
19	23	F	Yes, freq.	Yes, freq.	Nexus 7
20	24	M	Yes, freq.	Yes, freq.	Nexus 7
21	23	M	No	Yes, but few	Nexus 9
22	23	M	Yes, freq.	Yes, but few	Nexus 7
23	21	M	Yes, freq.	Yes, but few	Nexus 7
24	25	M	Yes, freq.	No	Nexus 7
25	21	F	Yes, but few	Yes, but few	Nexus 9
26	24	M	Yes, but few	No	Nexus 9
27	23	M	Yes, freq.	No	Nexus 9
28	23	M	Yes, but few	Yes, but few	Nexus 9
29	23	M	Yes, freq.	Yes, but few	Nexus 9
30	23	M	Yes, but few	No	Nexus 9

Table 3.3: Ratings for SUS statements, including positive statement total, negative statement total, SUS score and normalized SUS score (*Standard*), for each participant. Includes mean, median, mode, minimum and maximum.

P	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	+	-	Score	Normal
1	5	2	6	1	6	3	6	1	6	1	24	27	127.5	85.0
2	5	7	7	1	6	6	6	1	6	6	25	14	97.5	65.0
3	6	1	7	1	6	2	6	1	7	1	27	29	140	93.3
4	5	2	7	1	5	2	7	2	7	1	26	27	132.5	88.3
5	5	2	6	1	7	3	6	2	6	2	25	25	125	83.3
6	6	2	5	2	5	2	5	2	3	3	19	24	107.5	71.7
7	5	2	6	1	5	4	6	2	6	1	23	25	120	80.0
8	6	2	6	1	5	2	5	1	6	1	23	28	127.5	85.0
9	5	2	6	1	5	3	6	3	5	1	22	25	117.5	78.3
10	5	2	6	1	6	1	4	2	5	1	21	28	122.5	81.7
11	6	2	6	1	6	1	7	1	7	1	27	29	140	93.3
12	5	3	5	1	5	2	6	3	5	1	21	25	115	76.7
13	6	2	6	2	7	2	5	1	6	2	25	26	127.5	85.0
14	5	2	5	1	5	2	6	2	5	1	21	27	120	80.0
15	7	1	7	1	7	2	7	1	7	1	30	29	147.5	98.3
16	5	2	6	1	7	1	5	2	5	1	23	28	127.5	85.0
17	6	1	6	1	5	2	7	2	6	1	25	28	132.5	88.3
18	6	2	5	1	7	1	6	1	6	1	25	29	135	90.0
19	5	1	6	1	7	1	6	3	7	1	26	28	135	90.0
20	6	2	6	1	5	2	5	2	5	1	22	27	122.5	81.7
21	5	2	4	2	5	3	6	2	6	3	21	23	110	73.3
22	5	3	6	1	6	4	5	5	6	1	23	21	110	73.3
23	6	2	6	1	6	2	6	1	6	2	25	27	130	86.7
24	5	2	6	1	6	2	6	2	6	1	24	27	127.5	85.0
25	5	2	5	1	6	2	6	2	6	2	23	26	122.5	81.7
26	5	2	6	1	5	1	6	1	5	2	22	28	125	83.3
27	5	1	6	1	6	2	6	2	6	1	24	28	130	86.7
28	5	3	5	1	5	4	6	3	5	2	21	22	107.5	71.7
29	6	2	7	2	7	2	6	2	7	1	28	26	135	90.0
30	6	2	6	1	5	3	5	2	5	1	22	26	120	80.0
Mean	5.4	2.1	5.9	1.1	5.8	2.3	5.8	1.9	5.8	1.5	23.8	26.1	124.6	83.1
Med	5.0	2.0	6.0	1.0	6.0	2.0	6.0	2.0	6.0	1.0	23.5	27.0	126.3	84.2
Mode	5.0	2.0	6.0	1.0	5.0	2.0	6.0	2.0	6.0	1.0	25.0	28.0	127.5	85.0
Min	5.0	1.0	4.0	1.0	5.0	1.0	4.0	1.0	3.0	1.0	19.0	14.0	97.5	65.0
Max	7.0	7.0	7.0	2.0	7.0	6.0	7.0	3.0	7.0	6.0	30.0	29.0	147.5	98.3

Table 3.4: Ratings for SUS statements, including positive statement total, negative statement total, SUS score and normalized SUS score (*MovieWall*), for each participant. Includes mean, median, mode, minimum and maximum.

P	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	+	-	Score	Normal
1	3	5	3	1	4	3	2	5	5	5	12	6	45	30.0
2	6	1	6	1	6	1	6	6	6	1	25	25	125	83.3
3	6	2	6	1	6	1	5	1	7	1	25	29	135	90.0
4	4	5	5	1	5	3	4	2	5	3	18	21	97.5	65.0
5	3	6	3	2	6	7	2	5	4	2	13	13	65	43.3
6	2	6	3	3	5	3	3	5	3	4	11	14	62.5	41.7
7	5	3	5	1	6	2	5	3	5	2	21	24	112.5	75.0
8	3	6	3	1	3	1	6	6	2	1	12	20	80	53.3
9	5	1	7	1	6	2	6	2	5	2	24	27	127.5	85.0
10	3	7	4	1	3	7	2	6	2	4	9	10	47.5	31.7
11	3	4	4	3	2	3	4	6	3	3	11	16	67.5	45.0
12	2	3	4	1	2	4	3	7	3	2	9	18	67.5	45.0
13	2	6	4	4	5	5	2	7	4	4	12	9	52.5	35.0
14	3	2	5	3	2	2	3	5	3	3	11	20	77.5	51.7
15	3	2	4	1	5	1	4	4	5	1	16	26	105	70.0
16	3	6	4	1	5	2	3	6	3	2	13	18	77.5	51.7
17	4	3	6	1	2	4	5	3	4	1	16	23	97.5	65.0
18	5	4	5	2	3	1	5	4	5	2	18	22	100	66.7
19	4	4	4	1	5	3	3	5	4	1	15	21	90	60.0
20	6	5	5	1	5	3	6	2	6	1	23	23	115	76.7
21	2	6	6	3	3	5	3	5	3	6	12	10	55	36.7
22	3	6	6	1	5	6	4	5	7	2	20	15	87.5	58.3
23	3	4	4	3	3	4	4	5	5	2	14	17	77.5	51.7
24	3	4	4	1	5	3	4	5	4	2	15	20	87.5	58.3
25	4	4	4	1	4	4	3	4	4	3	14	19	82.5	55.0
26	3	3	5	1	5	3	5	3	4	2	17	23	100	66.7
27	3	3	5	2	6	3	4	4	5	2	18	21	97.5	65.0
28	5	4	5	1	3	2	5	5	4	1	17	22	97.5	65.0
29	3	6	4	2	4	3	3	5	3	6	13	15	70	46.7
30	4	5	5	2	4	3	3	5	5	1	15	18	82.5	55.0
Mean	3.6	4.2	4.6	1.6	4.3	3.1	3.9	4.5	4.3	2.4	15.6	18.8	86.2	57.4
Med	3.0	4.0	4.5	1.0	5.0	3.0	4.0	5.0	4.0	2.0	15.0	20.0	85.0	56.7
Mode	3.0	6.0	4.0	1.0	5.0	3.0	3.0	5.0	5.0	2.0	12.0	21.0	97.5	65.0
Min	2.0	1.0	3.0	1.0	2.0	1.0	2.0	1.0	2.0	1.0	9.0	6.0	45.0	30.0
Max	6.0	7.0	7.0	4.0	6.0	7.0	6.0	7.0	7.0	6.0	25.0	29.0	135.0	90.0

Table 3.5: Ratings and answers given for additional statements and questions seen in table 3.1, for each participant. Includes mean, median, mode, minimum and maximum when applicable.

P	A1	Q1	A2	A3	Q2	A4	A5
1	2	Yes	3	2	No		
2	7	Yes	7	7	Yes	7	7
3	3	Yes	3	5	No		6
4	2	Yes	2	4	No	4	3
5	2	Yes	5	2	No	4	2
6	2	Yes	2	2	No		3
7	3	Yes	5	3	No	3	4
8	2	Yes	6	2	Yes	1	2
9	6	Yes	6	6	No		2
10	2	Yes	3	1	Yes	1	2
11	2	Yes	4	2	Yes	5	2
12	1	Yes	2	1	No	4	2
13	1	Yes	2	2	No	4	5
14	2	Yes	1	2	No	4	3
15	1	Yes	2	1	No	3	1
16	2	Yes	1	1	No		
17	2	Yes	5	2	No	4	1
18	2	Yes	3	2	No		2
19	2	Yes	4	2	No	1	3
20	6	Yes	6	6	No	4	4
21	2	Yes	2	2	No	2	5
22	4	Yes	2	2	No	4	3
23	2	Yes	2	2	No	4	2
24	2	Yes	3	2	Yes	4	2
25	3	Yes	5	4	No	4	4
26	2	Yes	3	2	No	4	3
27	2	Yes	3	2	No	4	2
28	5	Yes	5	3	No		
29	2	Yes	2	2	Yes	2	2
30	3	Yes	4	3	Yes		4
Mean	2.6		3.4	2.6		3.5	3.0
Med	2.0		3.0	2.0		4.0	3.0
Mode	2.0		2.0	2.0		4.0	2.0
Min	1.0		1.0	1.0		1.0	1.0
Max	7.0		7.0	7.0		7.0	7.0

Table 3.6: User statements mentioned regarding the *Standard* interface, including frequency.

User statement	Times mentioned
Familiar	14
Better	10
More practical	6
More clear	4
Movies easier to foreground	4
Feels nicer	4
Has better sorting functionality	3
Direct	2
Communicates more	2
Chill	2
More personalization	1
More dynamic	1
Sorting of lists felt random	1
Add an option to modify list sorting	1
More control	1
Can see more than in Netflix	1
Easier to use	1
Easier to compare movies	1
Could be better	1
Should be more uniform	1

Table 3.7: User statements mentioned regarding the *MovieWall*, including frequency. Due to the number of different statements, they only shown if mentioned by participants more than once.

User statement	Frequency
Easily lost	16
Fun	11
Cool concept	11
Missing suggestions / personalization	11
Discovery / exploration focused	10
(Annoyed by) unclear connection between movies / clusters	9
No / bad structure, lack of structure	9
Tap functionality for headers	8
Missing dynamic rearrangement	7
Bad Implementation	7
Distracting	6
Hard to find movies	6
Overwhelming	5
Bad flow	5
Filters help	4
Exciting	4
Has a niche	4
More convenience features	3
Not practical	3
Has potential	3
Complementary	2
Automatic suggestions	2
Grows on you	2
Felt like just a novelty	2
More freedom	2

Table 3.8: Comparison between Nexus 9 and Nexus 7 values (based on data from table 3.2 and 3.3, grouping by device used as indicated in table 3.1).

Attribute	Mean		Median		Mode		Minimum		Maximum	
	N 9	N 7	N 9	N 7	N 9	N 7	N 9	N 7	N 9	N 7
S1 (S)	5.3	5.5	5.0	5.0	5.0	5.0	5.0	5.0	6.0	7.0
S2 (S)	2.3	1.9	2.0	2.0	2.0	2.0	1.0	1.0	7.0	3.0
S3 (S)	5.9	5.9	6.0	6.0	6.0	6.0	4.0	5.0	7.0	7.0
S4 (S)	1.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	2.0	1.0
S5 (S)	5.7	5.9	6.0	6.0	5.0	5.0	5.0	5.0	7.0	7.0
S6 (S)	2.7	1.9	2.0	2.0	2.0	2.0	1.0	1.0	6.0	4.0
S7 (S)	5.9	5.8	6.0	6.0	6.0	6.0	5.0	4.0	7.0	7.0
S8 (S)	1.7	2.1	2.0	2.0	2.0	2.0	1.0	1.0	3.0	5.0
S9 (S)	5.8	5.8	6.0	6.0	6.0	6.0	3.0	5.0	7.0	7.0
S10 (S)	1.9	1.1	2.0	1.0	1.0	1.0	1.0	1.0	6.0	2.0
+ (S)	23.7	23.9	24.0	23.0	25.0	23.0	19.0	21.0	28.0	30.0
- (S)	25.1	27.1	26.0	28.0	26.0	28.0	14.0	21.0	29.0	29.0
Total (S)	121.8	127.3	125.0	127.5	127.5	127.5	97.5	110.0	140.0	147.5
Normal (S)	81.2	84.9	83.3	85.0	85.0	85.0	65.0	73.3	93.3	98.3
S1 (M)	3.7	3.5	3.0	3.0	3.0	3.0	2.0	2.0	6.0	6.0
S2 (M)	4.3	4.1	5.0	4.0	6.0	4.0	1.0	1.0	6.0	7.0
S3 (M)	4.6	4.6	5.0	4.0	5.0	4.0	3.0	3.0	6.0	7.0
S4 (M)	1.7	1.5	1.0	1.0	1.0	1.0	1.0	1.0	4.0	3.0
S5 (M)	4.8	3.7	5.0	3.0	6.0	5.0	3.0	2.0	6.0	6.0
S6 (M)	3.2	3.1	3.0	3.0	3.0	3.0	1.0	1.0	7.0	7.0
S7 (M)	3.7	4.1	3.0	4.0	3.0	4.0	2.0	2.0	6.0	6.0
S8 (M)	4.3	4.7	5.0	5.0	5.0	5.0	1.0	2.0	7.0	7.0
S9 (M)	4.5	4.1	4.0	4.0	5.0	5.0	3.0	2.0	7.0	7.0
S10 (M)	2.9	1.9	2.0	2.0	1.0	2.0	1.0	1.0	6.0	4.0
+ (M)	16.2	15.1	15.0	15.0	12.0	9.0	11.0	9.0	25.0	24.0
- (M)	17.9	19.7	19.0	20.0	21.0	20.0	6.0	10.0	29.0	27.0
Total (M)	85.3	87.0	82.5	87.5	97.5	77.5	45.0	47.5	135.0	127.5
Normal (M)	56.9	58.0	55.0	58.3	65.0	51.7	30.0	31.7	90.0	85.0
A1	2.7	2.5	2.0	2.0	2.0	2.0	1.0	1.0	7.0	6.0
A2	3.5	3.3	3.0	3.0	2.0	2.0	2.0	1.0	7.0	6.0
A3	3.0	2.3	2.0	2.0	2.0	2.0	2.0	1.0	7.0	6.0
A4	3.8	3.3	4.0	4.0	4.0	4.0	2.0	1.0	7.0	5.0
A5	3.8	2.2	4.0	2.0	3.0	2.0	2.0	1.0	7.0	4.0

Table 3.9: All actions users can perform. Actions with the prefix "CS" can only be performed in the *Standard* interface. Actions with the prefix "CM" can only be performed in the *MovieWall* interface. Actions with the "C" prefix can be used in both.

Action	
CS1	Go Back
CS2	Home
CS3	More Like This
CS4	See All
CS5	See More
CS6	Start Drag
CS7	Start Drag (Vertical only)
CS8	Stop Drag
CS9	Use Filter
CM1	Start Touching MovieWall
CM2	Stop Touching MovieWall
CM3	Toggle Filter
C1	Add To WatchList
C2	Close Details Panel
C3	Open Details Panel
C4	Show Info
C5	Toggle WatchList

Table 3.10: Actions performed while using the *Standard* interface, per participant.

P	CS1	CS2	CS3	CS4	CS5	CS6	CS7	CS8	CS9	C1	C2	C3	C4	C5
1	7	0	2	0	0	77	15	92	5	5	13	22	0	6
2	2	0	1	1	2	80	3	83	0	0	3	3	0	0
3	7	2	3	1	0	114	29	143	5	12	19	27	1	12
4	2	0	0	1	1	26	4	30	0	0	4	4	0	0
5	3	1	1	5	1	57	20	77	4	1	7	7	1	4
6	5	1	1	3	2	139	21	159	0	4	5	4	0	4
7	1	2	4	0	0	100	30	130	2	1	14	14	0	4
8	2	2	1	1	0	151	33	183	1	1	9	22	1	6
9	5	1	0	0	1	167	43	208	4	0	12	11	0	4
10	3	1	1	1	1	28	13	41	1	3	7	8	1	3
11	2	1	2	1	1	90	18	104	0	1	3	5	0	2
12	9	2	3	1	2	106	20	124	2	3	11	9	1	15
13	6	2	1	2	1	138	11	147	4	10	15	22	0	10
14	4	0	1	1	2	41	16	57	0	1	6	6	0	6
15	2	2	2	0	2	28	9	35	3	2	8	8	0	4
16	3	0	0	1	2	295	46	337	0	0	9	24	0	0
17	4	0	0	2	2	37	12	48	0	0	4	4	0	2
18	3	2	2	0	0	90	13	103	4	8	12	13	1	3
19	6	1	1	2	2	72	22	93	2	1	9	8	0	2
20	3	1	1	0	1	353	63	415	1	2	14	14	1	14
21	0	1	1	2	1	86	13	98	0	1	12	12	1	1
22	3	0	1	0	0	60	21	81	3	1	10	11	0	0
23	3	2	6	2	1	71	28	98	1	2	13	13	1	2
24	4	0	5	2	0	79	19	97	2	3	16	16	1	3
25	2	1	2	1	1	79	15	94	0	4	14	13	1	1
26	0	0	1	0	0	193	3	196	0	10	5	14	0	0
27	2	1	5	0	0	83	16	99	3	2	16	20	1	1
28	2	2	5	0	0	47	14	61	1	4	13	14	0	4
29	3	1	6	0	0	91	26	117	1	7	22	21	0	2
30	7	0	4	2	4	99	33	129	1	4	8	8	0	2

Table 3.11: Actions performed while using the *MovieWall* interface, per participant.

P	CM1	CM2	CM3	C1	C2	C3	C4	C5
1	110	109	7	3	9	46	3	7
2	176	175	0	0	5	5	0	0
3	191	190	26	7	4	16	0	8
4	116	115	0	3	17	17	0	0
5	98	97	28	2	4	8	4	6
6	231	230	16	10	9	13	0	6
7	96	95	8	2	4	11	1	1
8	235	234	0	1	4	4	0	0
9	215	214	5	0	12	12	0	4
10	113	112	6	2	6	9	1	3
11	90	89	3	2	1	5	1	1
12	169	168	17	2	7	14	8	12
13	189	188	22	2	10	15	4	6
14	66	65	1	2	1	3	0	4
15	78	77	13	3	5	6	1	1
16	234	233	1	0	12	12	0	0
17	69	68	0	0	3	3	0	0
18	109	108	3	4	4	6	0	0
19	122	121	7	3	11	14	3	3
20	290	289	16	1	8	9	0	6
21	93	92	3	5	3	20	0	2
22	94	93	1	0	21	21	0	0
23	103	102	6	4	4	29	1	1
24	110	109	4	8	10	15	0	2
25	81	80	5	4	3	26	0	2
26	188	187	0	6	9	9	0	0
27	142	141	20	7	2	10	0	2
28	112	111	1	5	9	12	0	0
29	121	120	4	7	6	29	0	0
30	111	110	2	6	14	36	0	0

Limitations of the study and relevance for future work

From the results, it is clear that the current implementation of the *MovieWall* is severely lacking compared to the state-of-the-art, and requires future improvements towards usability and user experience. It is ill-suited as a replacement or even as a complementary function to existing interfaces. The major issues of the current implementation, including possible solutions, are described below.

Sixteen subjects of the user study got lost easily while using the *MovieWall*. Even though freedom during browsing was appreciated, a complete lack of guidance negatively impacted their user experience. Users frequently had to zoom out to regain their whereabouts, interrupting the flow of browsing. Future implementations should provide more information with which users can easily recognize where they are within the *MovieWall*. Yet, many options are dependent on the structure of new *MovieWall* iterations. As such, they can not be named blindly.

The structure of the current *MovieWall* was also frequently stated as a negative; Nine participants mentioned the structure was bad. Nine users also stated that connections between movies and clusters were either unclear, or lacking. Unlike the *Standard* interface, where users can alter the structure to suit their needs, subjects were stuck navigating a static grid. Additionally, the movies shown could sporadically change in genre, which can confuse participants. Future implementations should look into structures that make connections between clusters and movies clear. This may be done by emphasizing on a more dynamic approach, where a specific genre is highlighted while other genres are only considered as a secondary attribute for cluster placement.

Finally, new implementations should add options to personalize content; Eleven participants stated they clearly missed the recommendations and ability to personalize the movies on screen while using the *MovieWall*. They did have these options while using the *Standard* interface. New implementations of the *MovieWall* should allow users to change the content displayed by using their input. This input can then be used to find and bring movies of interest to the foreground or removing the filtered ones.

Our research has some weaknesses that may be handled differently in the

future. One issue is the focus on a small group of thirty participants, which primarily consisted of young adults, aged between 18 and 25 years old. Expanding to a wider, more culturally diverse and demographically varied audience may present a different outcome. Specifically, users who were familiar with state-of-the-art interfaces appeared to be slightly biased in favor of the *Standard* interface. While this bias may be an important obstacle for future implementations to overcome, it may also be less prevalent in a more varied audience.

Another weakness comes from the duration of our study. A short-term experiment is ill-suited for capturing long-term use of movie browsing interfaces. For one, favoritism towards the *Standard* interface may fade over time, and the same can be said for any novelty that the *MovieWall* interface brings. Furthermore, users did not have to watch any of the movies that they saw. A long-term study could emphasize more on users picking movies, watching them and measuring their satisfaction afterwards.

The last problem with our study is its emphasis on competition. Users did not only compare the two interfaces, but also rated one against the other. Moreover, the two interfaces were implemented separately from one another, with subjects unable to use both at the same time. As future implementations should look into the *MovieWall* as a complementary function, so too should research emphasize less on competition. One example may be comparing two interfaces, one with the *MovieWall* as a complementary function and one without it.

While our study has several weaknesses, it is unlikely that the strong preference for the *Standard* interface would disappear by performing a different comparative experiment. Furthermore, future usage of an application is often decided in the first five minutes. In this regard, it is clear that the *Standard* interface presented itself better than the *MovieWall*.

Given the results, we believe that the *MovieWall* concept has potential as an addition to state-of-the-art interfaces once the aforementioned issues are resolved. Until then, the benefits of the *MovieWall* implementation as a complementary function are small and uncertain compared to just using a state-of-the-art interface on its own.

References

[1] Ahn, H. J. A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem. *Information Sciences 178*, 1 (2008), 37–51.

[2] Athukorala, K., Głowacka, D., Jacucci, G., Oulasvirta, A., and Vreeken, J. Is exploratory search different? a comparison of information search behavior for exploratory and lookup tasks. *Journal of the Association for Information Science and Technology 67*, 11 (2016), 2635–2651.

[3] Brooke, J., et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry 189*, 194 (1996), 4–7.

[4] Cambazoglu, B. B., and Baeza-Yates, R. Scalability challenges in web search engines. In *Advanced topics in information retrieval*. Springer, 2011, pp. 27–50.

[5] Carvalhal, S., et al. Active exploration of online movie databases by visual navigation. Master's thesis, 2017.

[6] Chang, S.-F., Ma, W.-Y., and Smeulders, A. Recent advances and challenges of semantic image/video search. In *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on* (2007), vol. 4, IEEE, pp. IV–1205.

[7] Henzinger, M. R., Motwani, R., and Silverstein, C. Challenges in web search engines. In *ACM SIGIR Forum* (2002), vol. 36, ACM, pp. 11–22.

[8] Hunicke, R., LeBlanc, M., and Zubek, R. Mda: A formal approach to game design and game research. In *Proceedings of the AAAI Workshop on Challenges in Game AI* (2004), vol. 4, AAAI Press San Jose, CA, pp. 1–5.

[9] Kules, B., Capra, R., Banta, M., and Sierra, T. What do exploratory searchers look at in a faceted search interface? In *Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries* (2009), ACM, pp. 313–322.

[10] Lam, X. N., Vu, T., Le, T. D., and Duong, A. D. Addressing cold-start problem in recommendation systems. In *Proceedings of the 2nd international conference on Ubiquitous information management and communication* (2008), ACM, pp. 208–211.

[11] Low, T., Hentschel, C., Stober, S., Sack, H., and Nürnberger, A. Exploring large movie collections: Comparing visual berrypicking and traditional browsing. In *International Conference on Multimedia Modeling* (2017), Springer, pp. 198–208.

[12] Marchionini, G. Exploratory search: from finding to understanding. *Communications of the ACM 49*, 4 (2006), 41–46.

[13] Mirizzi, R., and Di Noia, T. From exploratory search to web search and back. In *Proceedings of the 3rd workshop on Ph. D. students in information and knowledge management* (2010), ACM, pp. 39–46.

[14] Nanopoulos, A., Rafailidis, D., Ruxanda, M. M., and Manolopoulos, Y. Music search engines: Specifications and challenges. *Information Processing & Management 45*, 3 (2009), 392–396.

[15] Nefkens, M. Moviewall: A novel interface concept for visual exploration of large movie collections. Master's thesis, 2017.

[16] Pang, P. C.-I., Verspoor, K., Chang, S., and Pearce, J. Conceptualising health information seeking behaviours and exploratory search: result of a qualitative study. *Health and Technology 5*, 1 (2015), 45–55.

[17] Pazzani, M. J., and Billsus, D. Content-based recommendation systems. In *The adaptive web*. Springer, 2007, pp. 325–341.

[18] Pereira, A. L. V., and Hruschka, E. R. Simultaneous co-clustering and learning to address the cold start problem in recommender systems. *Knowledge-Based Systems 82* (2015), 11–19.

[19] Ricci, F., Rokach, L., and Shapira, B. Recommender systems: introduction and challenges. In *Recommender systems handbook*. Springer, 2015, pp. 1–34.

[20] Schafer, J. B., Frankowski, D., Herlocker, J., and Sen, S. Collaborative filtering recommender systems. In *The adaptive web*. Springer, 2007, pp. 291–324.

[21] Thudt, A., Hinrichs, U., and Carpendale, S. The bohemian bookshelf: supporting serendipitous book discoveries through information visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2012), ACM, pp. 1461–1470.

[22] White, R. W., and Roth, R. A. Exploratory search: Beyond the query-response paradigm. *Synthesis lectures on information concepts, retrieval, and services 1*, 1 (2009), 1–98.

[23] Wildemuth, B. M., and Freund, L. Assigning search tasks designed to elicit exploratory search behaviors. In *Proceedings of the Symposium on Human-Computer Interaction and Information Retrieval* (2012), ACM, p. 4.