

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

A Service-oriented Perspective on the Summarization of Recommendations: Preliminary Experiment

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/1788454> since 2021-07-30T09:55:18Z

Publisher:

ACM

Published version:

DOI:10.1145/3450614.3464475

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

A Service-oriented Perspective on the Summarization of Recommendations: Preliminary Experiment

Noemi Mauro
University of Torino
Torino, Italy
noemi.mauro@unito.it

Liliana Ardissono
University of Torino
Torino, Italy
liliana.ardissono@unito.it

Zhongli Filippo Hu
University of Torino
Torino, Italy
zhonglifilippo.hu@unito.it

Gianmarco Izzi
University of Torino
Torino, Italy
gianmarco.izzi@unito.it

ABSTRACT

The explanation and justification of recommender systems' results are challenging research tasks. On the one hand, a model-based description that clarifies the reasoning approach behind the suggestions might be difficult to understand, or it might fail to convince the user, if (s)he does not agree on the applied inference mechanism. On the other hand, an aspect-based justification based on few characteristics might provide a partial view of items or, if more detailed, it might overload the user with too much information.

In order to address these issues, we propose a visual model aimed at justifying recommendations from a holistic perspective. Our model is based on a service-oriented summary of consumers' experience with items. We use the Service Journey Maps to extract data about the experience with services from online reviews, and to generate a visual summary of such feedback, based on evaluation dimensions that refer to all the stages of service fruition. Thanks to a graphical representation of these dimensions (based on bar graphs), and on the provision of on-demand data about the associated aspects of items, our model enables the user to overview the recommendation list and to quickly identify the subset of results that deserve to be inspected in detail for a final selection decision. A preliminary user study, based on the Apartment Monitoring application, has provided encouraging results about the usefulness and efficacy of our model to enhance user awareness and decision-making in the presence of medium-size recommendation lists.

CCS CONCEPTS

• **Information systems** → **Web searching and information discovery**; **Recommender systems**.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

UMAP '21 Adjunct, June 21–25, 2021, Utrecht, Netherlands

© 2021 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-8367-7/21/06...\$15.00

<https://doi.org/10.1145/3450614.3464475>

KEYWORDS

review-based recommender systems, summarization of recommendation lists, justification or results, Service Journey Maps

ACM Reference Format:

Noemi Mauro, Zhongli Filippo Hu, Liliana Ardissono, and Gianmarco Izzi. 2021. A Service-oriented Perspective on the Summarization of Recommendations: Preliminary Experiment. In *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '21 Adjunct)*, June 21–25, 2021, Utrecht, Netherlands. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3450614.3464475>

1 INTRODUCTION

The ranked list of items presented by e-commerce platforms as a result of an information search task poorly helps content interpretation because it does not justify the returned results. The main obstacles to users' decision-making concern understanding why items are relevant to the submitted query, and how they are related to each other [19]. In order to increase user awareness and to enhance the transparency of recommendations, some systems attempt to explain how the system produced them, e.g., by describing degree or type of match between user and items [14, 29, 37]. Differently, other systems, such as [26–28], justify recommendations, regardless of how they have been obtained, by extracting the main aspects about items from consumer feedback.

Both approaches have limitations. On the one hand, explaining why an item is suggested might fail to help the user's decision-making because (s)he disagrees with the rationale behind the suggestions. On the other hand, as review summarization conveys information about the experience with items, it is useful to highlight pros and cons of products and services [3, 12]. However, it generates descriptions that make direct reference to specific item features; e.g., see [7, 31, 41]. Thus, it might produce either partial views on items (if a few aspects are mentioned), or detailed presentations that hinder the efficiency in the exploration of medium-size suggestion lists.

Another key issue is that, as specified by service design models like the Service Journey Maps [34, 35], the experience with items can involve different stages of interaction with the provider, all of which impact on consumer satisfaction. In other words, items have to be described by taking the overall context of fruition into account. For instance, focusing on the e-commerce domain, the user experience starts with the search for products on the website and

it ends with their delivery to the customer. We claim that, in order to enhance user awareness, all these phases should be considered both in the evaluation and in the presentation of results to the user.

In this work, we focus on the presentation of results. According to marketing research [13], product comparison is a key step in the selection of goods. We thus aim at helping the user to efficiently identify, within the suggestion list, a small set of preferred options that (s)he can easily compare, from a holistic viewpoint on item fruition. For this purpose, we propose a novel, visual summarization model, which reflects customers' perspective by generating a service-oriented overview of a recommendation list. We employ the Service Journey Maps [34, 35] to define high-level evaluation dimensions of a service associated with the stages of fruition by its users. Different from the visual approach presented in [7], which summarizes quantitative and qualitative data about specific item features, our model employs bar graphs at a more abstract level to represent the satisfaction degree reported by previous users with respect to such evaluation dimensions. Moreover, it enables the user to receive a qualitative justification of the evaluations based on the aspects of user experience emerging from consumer feedback.

Our work is framed in the Apartment Monitoring application [22], which supports the exploration of homes by analyzing the online reviews collected by the Airbnb website (<https://airbnb.com>). This system enables the user to search for homes and to analyze reviews by focusing on different dimensions of the overall service experience, such as the appreciation of the host (i.e., of the person who rents the home) by guests, and the in-apartment experience. In this work, we are interested in understanding whether a visual summary of the evaluations concerning the stages of interaction with items helps users quickly select the preferred options, or whether additional data is needed for decision-making.

We tested our model in a preliminary user study that involved 11 people. We found that, while participants wanted to see complete information about a home before booking it, they perceived our model as particularly helpful to quickly identify a small set of options which deserve to be further inspected, within the list of recommended ones. These findings encourage the integration of a visual, service-based summarization model within an online catalog to support user awareness about the available items and to help decision-making.

Section 2 positions our work within the related one and gives some background notions to the reader. Section 3 presents the Apartment Monitoring application. Section 4 describes our visual model. Section 5 describes the user study that we carried out and discusses the results we obtained. Section 6 concludes the paper and outlines our future work.

2 BACKGROUND AND RELATED WORK

2.1 Service Journey Maps

The Service Journey Maps (SJMs) [34, 35] have been introduced to support the design and development of physical and online products and services by focusing on the way customers experience them. A SJM is a visual description of the user experience with a service, such as a hotel, or an online retailer, which models the various stages a person encounters during service fruition. The graphic visualization follows a temporal line from the start point (e.g., enter

website or enter shop) to the end one (e.g., on site or online customer care) in order to highlight and understand the various stages a person gets in touch with to benefit from the service; these stages are named *touchpoints*. A SJM thus provides a holistic view of the context in which products and services can be evaluated.

We use SJMs to analyze and organize the consumer feedback collected by Airbnb. Starting from the identification of the touchpoints of this service, we derive the high-level evaluation dimensions to be used in the presentation of homes. Then, we extract from online reviews the aspects and opinions concerning such dimensions. In this way, we can provide the user with quantitative and qualitative information about previous consumers' perceptions.

2.2 Explaining/Justifying Recommendations

The explanation of results is a mainstream topic in recommender systems research because it enhances both the transparency of systems [37], and users' trust in the generated solutions [2]. Usually, explanations are based on the algorithm behind the recommendation process [14, 29, 37]. Some aspect-based recommender systems explain suggestions by highlighting the features of items which match, or mismatch, the target user's preferences [25]. Other ones support feature-based item comparison [6, 7, 23, 31], or information exploration based on the visualization of the relevance of items with respect to the keywords of search queries [5, 9]. Graph-based recommenders use the graph that connects users and items as explanations [1, 40], or as justifications of suggestions [25–27]. Finally, some recommenders pursue the idea of explaining while recommending; e.g., see [10, 21, 24].

In hybrid recommender systems, the presentation of results aims at explaining suggestions under multiple perspectives of relevance. MyMovieFinder [20] separately shows the recommenders that support a suggested item, while RelevanceTuner [38] uses stackable bars to integrate this type of information into a compact view. TalkExplorer [39] and IntersectionExplorer [4] use bidimensional graphs, or grid layouts, to show multiple dimensions of relevance. Moreover, Venn diagrams are used to overview suggestions [16] and they are combined with color bars to identify the recommender systems which contribute to the suggestions [30]. Finally, HyPER [17] explains the sources of recommendations within a probabilistic logic framework.

Our work justifies recommendations by generating a visual, holistic summary of consumer perceptions about items. The classical aspect-based justification (e.g., [7, 25, 31]) focuses on few peculiarities of items, hardly representing a complete view about the elements of the suggestion list. Differently, by employing service touchpoints as summarization pivots, we provide a comprehensive, high-level evaluation of user experience, which can be analyzed in depth by selecting the evaluation dimensions of interest. This incremental access to information makes it possible to deal with medium-size recommendation lists.

3 APARTMENT MONITORING

Apartment Monitoring supports the exploration of homes by retrieving their details and reviews from the Airbnb website.¹ The

¹ Airbnb data can be downloaded from <http://insideairbnb.com/get-the-data.html>, under Creative Commons CC0 1.0 Universal (CC0 1.0) "Public Domain Dedication"

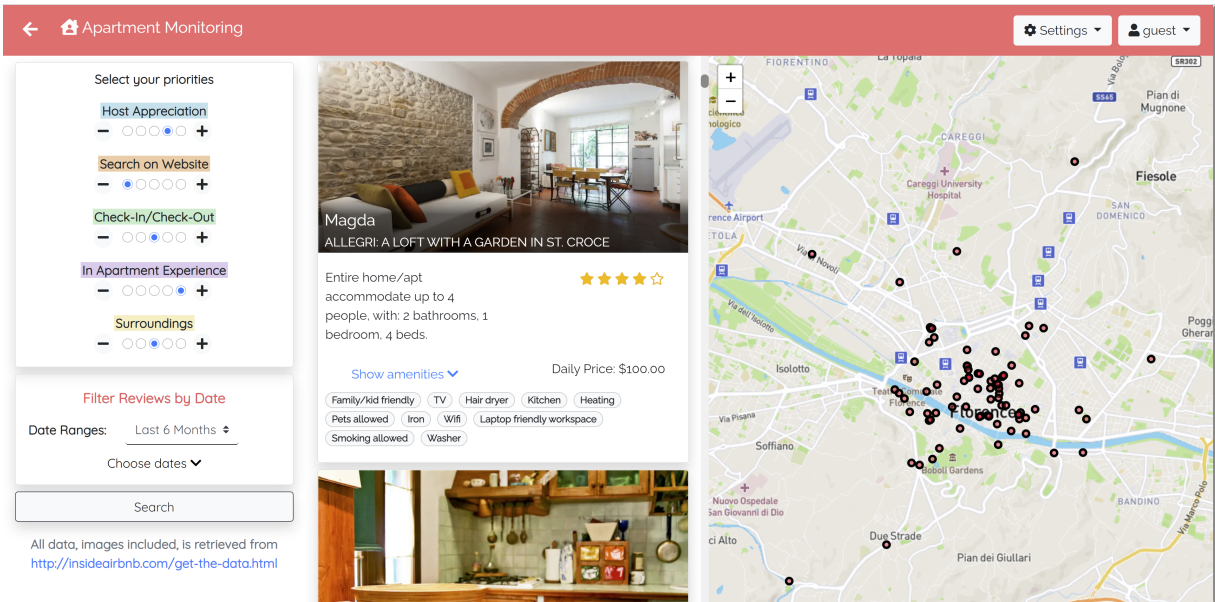


Figure 1: Apartment Monitoring: visualization of recommended homes. Data is from Inside Airbnb.

system personalizes the recommendation of items in a multi-criteria approach based on evaluation dimensions that are associated with the stages of interaction with the overall service, from the home-booking phase to the final check-out. See Sections 3.1 and 3.2.

Figure 1 shows a portion of the user interface of Apartment Monitoring. The left sidebar displays the form for the specification of the user profile (currently elicited from the user) and a temporal filter to select the online reviews to be used for recommendation. The user profile is based on the evaluation dimensions defined in Section 3.1. User preferences, denoted as priorities over evaluation dimensions, are expressed in the [1, 5] scale. The right portion of the user interface shows the scrollable ranked list of suggested homes and their location in a geographical map. For each home, the application shows the name of the host and of the home, a picture, a short description, the daily price and the list of offered amenities. Moreover, it shows the estimated rating of the home, expressed as a star list in [1, 5]. By clicking on a home, the user receives detailed information about it, including the reviews it received.

3.1 Service-based Representation of Homes

As described in detail in [22], the specification of a domain model supporting the holistic presentation of consumer opinions about a service is based on the following tasks:

- The specification of the main stages of interaction with the service, before, during and after service fruition.
- The identification of the evaluation dimensions associated with each stage, based on the aspects of the service emerging from consumer feedback.

For the first task, we designed a Service Journey Map describing the stages of a typical home renting experience from the viewpoint of

license. The Inside Airbnb service provides home details and reviews but it does not provide the ratings given by users to the homes.

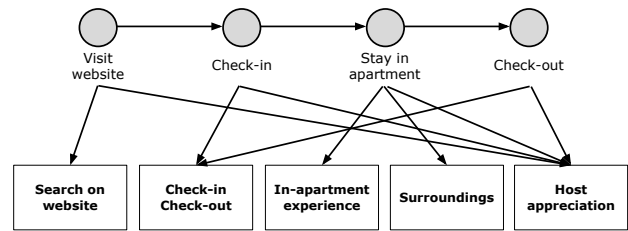


Figure 2: Service Journey Map of the home booking experience (represented as a workflow) and evaluation dimensions associated to its stages (depicted as rectangles).

the guest, i.e., ignoring the backstage activities that are not visible to her/him. For the definition of the map, we got inspiration from standard Service Journey Maps available for hotel booking and we adapted them to the home booking domain. This led to the definition of the workflow shown in the upper portion of Figure 2.

For the identification of the evaluation dimensions associated with each stage of the Service Journey Map (lower portion of the figure), we took into account the studies carried out in the literature about home and hotel booking [8, 33, 42]. Moreover, by extracting the aspects that occur in the analyzed Airbnb reviews, we noticed that (i) *Check-in* and *Check-out* are usually coupled in reviews and they are associated to the same terms; (ii) *Stay in apartment* can be associated to two different dimensions because the aspects related to the apartment interiors (furniture, comfort, services) are frequently separated in the reviews from those concerning its surroundings, e.g., view, shops, and so forth. We also noticed that reviewers frequently distinguish their evaluation of the host from that of the home, and that the interaction with the host, and her/his properties,

represent a relevant evaluation dimension crossing all the service stages. We thus defined the following high-level dimensions for the description of homes:

- *Host appreciation*: this represents guests' perceptions of the host and of the interaction with her/him at any time of service fruition.
- *Search on website*: this refers to the perceived effort in looking for information about the home in the Airbnb website. We overlook this dimension that concerns the Airbnb platform.
- *Check-in/Check-out*: this summarizes guests' experience at check-in and check-out times, e.g., timeliness.
- *In apartment experience*: this represents guests' perceptions within the apartment, e.g., cleanliness and comfort.
- *Surroundings*: this describes the perception of the area where the home is located, in terms of aspects such as available services, quietness, and so forth.

For each dimension, we built a thesaurus of the associated aspects. This is used to support the extraction of aspects, and the consequent reference to evaluation dimensions, in the sentences of the reviews.

3.2 Personalized Recommendation of Homes

The recommendation of items is out of the scope of this paper but we provide a short description of it for completeness.

The application ranks homes for their presentation on the basis of the sentiment emerging from online reviews, and it selects the best ranked items for presentation in the user interface. The score of each home h is estimated in two steps:

- (1) The application infers the sentiment about the evaluation dimensions defined in Section 3.1 and it normalizes it in the $[1, 5]$ scale. The sentiment about an evaluation dimension D is computed by applying sentiment analysis to the aspects associated to D that occur in the reviews of h . The thesaurus of D is used for this purpose.
- (2) The score of h is estimated as the weighted mean of the sentiment computed in step 1, using the priorities expressed in the user profile as weights. Thus, strong preferences have more impact on the estimation of the score than weak ones.

4 PROPOSED VISUALIZATION MODEL

The star-based rating of a home h shown in Figure 1 provides a coarse-grained view of its estimated suitability for the user. Within a recommendation list that includes more than a few options, this visualization can challenge decision-making because, when comparing homes to each other, the user should inspect their details in a possibly lengthy analysis task. In order to mitigate this issue, we propose an interactive model that summarizes both quantitative and qualitative information about previous guests' experience with homes, inferred from their reviews.

Concerning **quantitative information**, we propose a graphical view that assesses previous guests' perceptions of the booking experience with respect to the evaluation dimensions described in Section 3.1. Figure 3 shows a portion of the new user interface of Apartment Monitoring: the star-based score of each home is replaced with a bar graph that represents the emerging sentiment for each evaluation dimension. The values of the bars are obtained

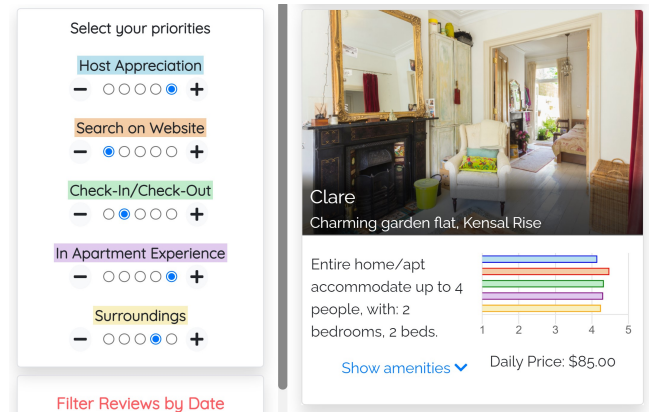


Figure 3: Visualization of a home with quantitative summary of consumer feedback. The amenities are hidden. Data is from Inside Airbnb.

by applying the method described in step 1 of Section 3.2. We use color coding to associate the bars with the priorities reported in the user profile section, which plays the role of a legend while the user scrolls the home list.

The presentation of **qualitative information** about homes is under development and is sketched in Figure 4. Our idea is that of supporting the evaluations conveyed by the bars of the graph with the related aspects of each home emerging from its reviews. We want that the user is able to select the homes for further analysis without being overloaded by information. Therefore, qualitative data will be grouped by evaluation dimension and it will be visualized on demand, by clicking on the bars of the graph. For each bar, and corresponding evaluation dimension D , we plan to visualize a list of relevant aspects which characterize a home h from the viewpoint of D . These are the aspects occurring in the reviews of h that belong to the thesaurus of D . As a first relevance measure, we will use the frequency of terms in the reviews.

5 PRELIMINARY EXPERIMENT

We investigate whether the summarization of feedback about homes provided by our model, and the incremental presentation of information, based on the visualization of quantitative data and of on-demand qualitative data about homes, enhance decision-making. In detail, we are interested in answering the following research questions:

- RQ1: Does the bar graph of a home h give users a quick assessment of previous guests' perceptions emerging from the reviews about h ?
- RQ2: Does the overall visual model (bar graphs + qualitative data consisting of aspects) help users focus on the most promising homes within a recommendation list?
- RQ3: Is the proposed visual model (bar graphs + qualitative data) enough for users to select the home they would like to book? Or do users need to view more home details, such as amenities, reviews, etc., to make a decision?

We carried out a preliminary experiment to shed light on these questions and steer the development of our visual model. In this

experiment, the bar graphs are generated by the application, as shown in Figure 3, while the on-demand aspects of homes are handcrafted.

5.1 Experiment Setup

We conducted a user study by exploiting a small portion of Apartment Monitoring to test whether the service-oriented information provided by our visual model enhances item filtering and evaluation. By showing the user a set of homes, the experiment investigates the level of support to the assessment of their suitability for her/him.

As home-booking is a high-investment domain², we assume that the user needs detailed information about items, such as the offered amenities, to make a renting decision. However, we aim at understanding whether (s)he can identify a small set of promising homes to be inspected in detail by looking at the visual summary described in Section 4. Therefore, we asked participants to evaluate whether the homes we proposed were *good candidates* for a renting decision, or not.

Figure 4 shows the test application we used for the study. The user interface omits several details about homes. Indeed, we decided to hide amenities and pictures to force participants make their decisions on the sole basis of the quantitative/qualitative data provided by our model. Specifically, as noticed in previous research about the evaluation of explanations [37], pictures and other types of information might influence rating behavior. We defined two tasks in order to test two versions of the visual model:

- TASK1: For each presented home, we only showed the bar graph representing the *Host appreciation*, *Check-in/Check-out*, *In apartment experience* and *Surroundings* evaluation dimensions.
- TASK2: For each presented home, we linked every bar of the graph to a set of aspects that justify its value. The user could view the aspects concerning a specific evaluation dimension by clicking on the associated bar.

We asked participants to rate 10 homes (5 for each task). The test application enabled them to specify that they did not know which rating to give, i.e., to opt out. This is useful to understand whether both versions of the model support decision-making.

The study was a within-subjects design one. We considered each treatment condition as an independent variable and every person received both treatments. We counterbalanced the order of tasks to minimize the impact of result biases and the effects of practice and fatigue. People joined the user study on a voluntary basis, without any compensation, and they gave their informed consent to participate in the study. The user study took place live, in video calls with shared screen due to COVID-19 pandemic. In other words, we did not use any platforms such as Amazon Mechanical Turk or similar to recruit people, or to manage the tasks of the study.

One person at a time participated in the study, which lasted about 15 minutes. We allowed people to take as much time as they needed to complete the experiment, in order to comply with diverse backgrounds and levels of confidence with technology. The study was organized as follows:

²Similar to [18, 36], the definition of “investment” rests on the concept of price. For example, buying a movie is a low investment, while home-booking is a high investment.

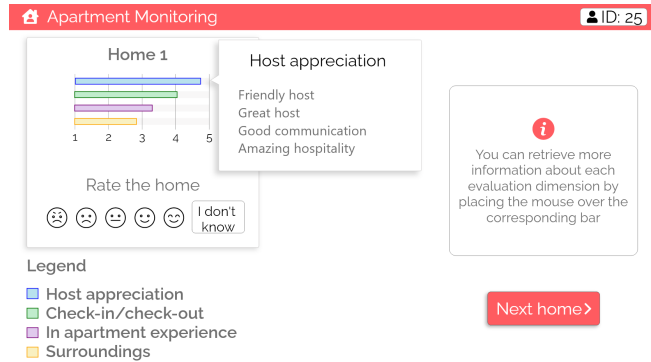


Figure 4: User interface of our test application. The *Host appreciation* bar has been expanded to view qualitative information about the host.

- (1) We asked the participant to answer a pre-test questionnaire designed to assess demographic information, cultural background, and familiarity with booking platforms.
- (2) While (s)he performed the tasks, the experimenter took notes about her/his voice comments using the think aloud protocol [15]. We did not put any time restrictions on question answering.
- (3) After the completion of each task, the participant filled in a post-task questionnaire to evaluate the version of the model (s)he used. This questionnaire was aimed at understanding whether the data provided by our visual model helped her/him to select good candidates for booking, or not.

5.2 Results and Discussion

5.2.1 Demographic Data and Background. We recruited 11 participants (54.54% women; 45.46% men, 0% other, 0% not declared). They are between 19 and 57 years old, with a mean value of 36.36. They are part of the university staff (researchers and professors) and students, as well as people working in the industry. In the pre-test questionnaire, we analyzed their background and familiarity with technology. Regarding the education level, 27.28% attended the high school, 36.36% the university, 18.18% have a Ph.D, and 18.18% attended the middle school. All of them regularly use the Internet.

We also investigated participants’ trust in other actors, taking inspiration from the ResQue questionnaire for recommender systems [32]. We asked people if they tend to trust a person or thing, even though they have little knowledge about it. Results are as follows: 9.09% of participants declared that they very probably trust it, 18.18% probably trust it, 54.55% probably do not trust it, and 18.18% very probably do not trust it. We can thus hypothesize that the majority would desire to have a relevant amount of data to evaluate the homes proposed by the system, rather than totally depend on black-box recommendations. This is important to assess the awareness support provided by our model.

5.2.2 Task Results. In TASK1, 12.73% of the ratings were “I don’t know”; in TASK2, participants rated all the homes. This means that, if the bar graph does not provide any qualitative data to justify the

values of the evaluation dimensions, it poorly supports decision-making. Notice that, in the post-task questionnaire after completing TASK1, 54.54% of people declared that the given information was not enough to rate the homes. Differently, in TASK2, the number of people encountering decision-making problems was lower. Specifically, 27.27% said that they would have preferred to have more data about the features of homes, such as the number of rooms, and the offered amenities. Only one person declared that (s)he was not able to understand the information conveyed by the bar graph in TASK1. From the think-aloud protocol we collected various comments, expressed in the Italian language, which we summarize:

- (1) Two people declared that, in a real word application, the bar graph is useful to filter out a home that does not deserve to be further analyzed because it performs badly on an evaluation dimension they care about. For instance, a participant explained that (s)he discarded a home by only looking at its graph, because the host was badly evaluated and (s)he did not intend to interact with difficult people.
- (2) Some participants were very strict on the evaluation dimensions that they considered important. For example, three people stated that, if the *Host appreciation* had a low value, they would not consider the other dimensions to rate the home. This is strictly related to finding (1) above.
- (3) Three people did not care much about the specific values of the bars in the graph. Differently, they compared the bars to each other in order to see on which dimensions a home was rated better or worse by previous guests.
- (4) For what concerns the qualitative data about evaluation dimensions (i.e., the aspects presented on-demand), four people said that they are useful to implicitly "tune" the values of the bars. For instance, suppose that a home h receives a low evaluation regarding a dimension d , and that the justification of d is based on aspects that are irrelevant to the user. Then, (s)he might implicitly increase the evaluation of h .
- (5) A participant suggested to apply color coding to the visualization of aspects. For instance, similar to [11], (s)he proposed to highlight the good and bad aspects of an individual home in green or red, respectively. Moreover, (s)he said that (s)he would have liked to filter aspects to quickly assess whether a specific one, such as WiFi, was good or not in a home.
- (6) A participant stated that qualitative data is more important than the bar graph because it helps you to select the rating to be given to a home. That person also said that (s)he gave more importance to the justifications presented by the system than to the bar graph because, usually, people put more effort in writing than in rating.
- (7) Not surprisingly, a participant said that the information provided by our model is very useful to select the homes that (s)he would like to analyze, but that (s)he would not feel confident in booking a home on the sole basis of this data. In other words, our visual model supports information filtering. However, more details are needed to make high-cost decisions such as renting a home.

The findings of this study are encouraging: they show that our model enables the user to efficiently reduce the solution space according to her/his preferences. While this aspect is particularly

relevant to information exploration research, it contributes to improve recommender systems, as well. In fact, by proposing relatively long recommendation lists and, at the same time, supporting the user in an active selection of the most relevant items, the system can increase the variety in results. Moreover, it can at least partially address the trade-off between accuracy and diversity. These findings also highlight that the quantitative data provided by bar graphs has to be complemented with qualitative data about items in order to support information filtering in an efficacious way, because qualitative data helps the interpretation of quantitative information.

6 CONCLUSIONS AND FUTURE WORK

We described a novel visualization approach, based on Service Journey Maps, for the justification of item recommendations. Our model supports the exploration of consumer feedback from a service-oriented viewpoint. It generates interactive overviews of recommendation lists by presenting a quantitative summary, enriched with on-demand, qualitative data about previous users' experience with items. We applied our model to the home-booking domain, within the Apartment Monitoring application. A preliminary user study has shown that the model successfully helps users in filtering the information space because they receive an effective, holistic overview of consumer feedback. Moreover, the on-demand qualitative data about previous consumers' experience enhances user awareness about items.

Our future work includes various research tasks. First of all, we have to complete the development of our visual model. We plan to develop automatic strategies, e.g., based on different relevance metrics [26], to select the aspects that should be visualized as qualitative data, out of the large set that could be extracted from the reviews. Moreover, we are interested in testing different approaches for the graphical presentation of quantitative and qualitative data about items. We also plan to conduct a larger user study to retrieve more extensive experimental results. Furthermore, we plan to instantiate our model on a different application domain (e.g., the sales of experience products) to assess its applicability to heterogeneous items. The specification of a new application domain can be facilitated by the existence of Service Journey Maps defined for various services, which can be adapted to the peculiarities of the selected domain.

ACKNOWLEDGMENTS

This work has been funded by the University of Torino under grant ARDL_RILO_2019.

REFERENCES

- [1] Saeed Amal, Chun-Hua Tsai, Peter Brusilovsky, Tsvi Kuflik, and Einat Minkov. 2019. Relational social recommendation: application to the academic domain. *Expert Systems with Applications* 124 (2019), 182–195. <https://doi.org/10.1016/j.eswa.2019.01.061>
- [2] Shlomo Berkovsky, Ronnie Taib, and Dan Conway. 2017. How to Recommend? User Trust Factors in Movie Recommender Systems. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces (Limassol, Cyprus) (IUI '17)*. Association for Computing Machinery, New York, NY, USA, 287–300. <https://doi.org/10.1145/3025171.3025209>
- [3] Eda Bilici and Yücel Saygın. 2017. Why do people (not) like me?: Mining opinion influencing factors from reviews. *Expert Systems with Applications* 68 (2017), 185–195. <https://doi.org/10.1016/j.eswa.2016.10.001>
- [4] Bruno Cardoso, Gayane Sedrakyan, Francisco Gutiérrez, Denis Parra, Peter Brusilovsky, and Katrien Verbert. 2019. IntersectionExplorer, a multi-perspective

- approach for exploring recommendations. *International Journal of Human-Computer Studies* 121 (2019), 73 – 92. <https://doi.org/10.1016/j.ijhcs.2018.04.008>
- [5] Joseph Chee Chang, Nathan Hahn, Adam Perer, and Aniket Kittur. 2019. Search-Lens: composing and capturing complex user interests for exploratory search. In *Proceedings of the 24th International Conference on Intelligent User Interfaces* (Marina del Ray, California) (IUI '19). ACM, New York, NY, USA, 498–509. <https://doi.org/10.1145/3301275.3302321>
 - [6] Li Chen and Feng Wang. 2017. Explaining recommendations based on feature sentiments in product reviews. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces* (Limassol, Cyprus) (IUI '17). Association for Computing Machinery, New York, NY, USA, 17–28. <https://doi.org/10.1145/3025171.3025173>
 - [7] Li Chen, Feng Wang, Luole Qi, and Fengfeng Liang. 2014. Experiment on sentiment embedded comparison interface. *Knowledge-Based Systems* 64 (2014), 44–58. <https://doi.org/10.1016/j.knsys.2014.03.020>
 - [8] Mingming Cheng and Xin Jin. 2019. What do Airbnb users care about? An analysis of online review comments. *International Journal of Hospitality Management* 76 (2019), 58 – 70. <https://doi.org/10.1016/j.ijhm.2018.04.004>
 - [9] Cecilia di Sciascio, Vedran Sabol, and Eduardo E. Veas. 2016. Rank As You Go: User-Driven Exploration of Search Results. In *Proceedings of the 21st International Conference on Intelligent User Interfaces* (Sonoma, California, USA) (IUI '16). Association for Computing Machinery, New York, NY, USA, 118–129. <https://doi.org/10.1145/2856767.2856797>
 - [10] Ruihai Dong and Barry Smyth. 2017. User-based Opinion-based Recommendation. In *Proceedings 26th IJCAI*. Melbourne, Australia, 4821–4825.
 - [11] Fatih Gedikli, Dietmar Jannach, and Mouzhi Ge. 2014. How should I explain? A comparison of different explanation types for recommender systems. *International Journal of Human-Computer Studies* 72, 4 (2014), 367–382. <https://doi.org/10.1016/j.ijhcs.2013.12.007>
 - [12] Anindya Ghose and Panagiotis G. Ipeirotis. 2011. Estimating the helpfulness and economic impact of product reviews: mining text and reviewer characteristics. *IEEE Transactions on Knowledge and Data Engineering* 23, 10 (2011), 1498–1512. <https://doi.org/10.1109/TKDE.2010.188>
 - [13] Gerald H aubl and Valerie Trifts. 2000. Consumer Decision Making in Online Shopping Environments: The Effects of Interactive Decision Aids. *Marketing Science* 19, 1 (2000), 4–21. <https://doi.org/10.1287/mksc.19.1.4.15178>
 - [14] Jonathan L. Herlocker, Joseph A. Konstan, and John Riedl. 2000. Explaining Collaborative Filtering recommendations. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work* (Philadelphia, Pennsylvania, USA) (CSCW '00). Association for Computing Machinery, New York, NY, USA, 241–250. <https://doi.org/10.1145/358916.358995>
 - [15] Anker Helms Jørgensen. 1990. Thinking-aloud in user interface design: a method promoting cognitive ergonomics. *Ergonomics* 33, 4 (1990), 501–507. <https://doi.org/10.1080/00140139008927157>
 - [16] Pigi Kouki, James Schaffer, Jay Pujara, John O'Donovan, and Lise Getoor. 2017. User preferences for hybrid explanations. In *Proceedings of the Eleventh ACM Conference on Recommender Systems* (Como, Italy) (RecSys '17). Association for Computing Machinery, New York, NY, USA, 84–88. <https://doi.org/10.1145/3109859.3109915>
 - [17] Pigi Kouki, James Schaffer, Jay Pujara, John O'Donovan, and Lise Getoor. 2019. Personalized explanations for hybrid recommender systems. In *Proceedings of the 24th International Conference on Intelligent User Interfaces* (Marina del Ray, California) (IUI '19). Association for Computing Machinery, New York, NY, USA, 379–390. <https://doi.org/10.1145/3301275.3302306>
 - [18] David N. Laband. 1991. An Objective Measure of Search versus Experience Goods. *Economic Inquiry* 29, 3 (1991), 497–509. <https://doi.org/10.1111/j.1465-7295.1991.tb00842.x>
 - [19] Jimmy Lin, Michael DiCuccio, Vahan Grigoryan, and W. John Wilbur. 2008. Navigating information spaces: a case study of related article search in PubMed. *Information Processing & Management* 44, 5 (2008), 1771 – 1783. <https://doi.org/10.1016/j.ipm.2008.04.002>
 - [20] Benedikt Loepp, Katja Herrmann, and Jürgen Ziegler. 2015. Blended recommending: integrating interactive information filtering and algorithmic recommender techniques. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) (CHI '15). ACM, New York, NY, USA, 975–984. <https://doi.org/10.1145/2702123.2702496>
 - [21] Yichao Lu, Ruihai Dong, and Barry Smyth. 2018. Why I like It: Multi-Task Learning for Recommendation and Explanation. In *Proceedings of the 12th ACM Conference on Recommender Systems* (Vancouver, British Columbia, Canada) (RecSys '18). Association for Computing Machinery, New York, NY, USA, 4–12. <https://doi.org/10.1145/3240323.3240365>
 - [22] Noemi Mauro, Liliana Ardissono, Sara Capecchi, and Rosario Galioto. 2020. Service-Aware Interactive Presentation of Items for Decision-Making. *Applied Sciences, Special Issue Implicit and Explicit Human-Computer Interaction* 10, 16 (2020), 5599. <https://doi.org/10.3390/app10165599>
 - [23] Martijn Millecamp, Nyi Nyi Htun, Cristina Conati, and Katrien Verbert. 2020. What's in a User? Towards Personalising Transparency for Music Recommender Interfaces. In *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (Genoa, Italy) (UMAP '20). Association for Computing Machinery, New York, NY, USA, 173–182. <https://doi.org/10.1145/3340631.3394844>
 - [24] K. Muhammad, A. Lawlor, E. Rafter, and B. Smyth. 2015. Great Explanations: opinionated explanations for recommendations. In *Case-Based Reasoning Research and Development*, Eyke Hüllermeier and Mirjam Minor (Eds.). Springer International Publishing, Cham, 244–258. https://doi.org/10.1007/978-3-319-24586-7_17
 - [25] Khalil Ibrahim Muhammad, Aonghus Lawlor, and Barry Smyth. 2016. A Live-User Study of Opinionated Explanations for Recommender Systems. In *Proceedings of the 21st International Conference on Intelligent User Interfaces* (Sonoma, California, USA) (IUI '16). Association for Computing Machinery, New York, NY, USA, 256–260. <https://doi.org/10.1145/2856767.2856813>
 - [26] Cataldo Musto, Marco de Gemmis, Pasquale Lops, and Giovanni Semeraro. 2020. Generating post hoc review-based natural language justifications for recommender systems. *User-Modeling and User-Adapted Interaction* 27 (2020). <https://doi.org/10.1007/s11257-020-09270-8>
 - [27] Cataldo Musto, Fedelucio Narducci, Pasquale Lops, Marco de Gemmis, and Giovanni Semeraro. 2019. Linked open data-based explanations for transparent recommender systems. *International Journal of Human-Computer Studies* 121 (2019), 93 – 107. <https://doi.org/10.1016/j.ijhcs.2018.03.003> Advances in Computer-Human Interaction for Recommender Systems.
 - [28] Jianmo Ni, Jiacheng Li, and Julian McAuley. 2019. Justifying recommendations using distantly-labeled reviews and fine-grained aspects. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. Association for Computational Linguistics, Hong Kong, China, 188–197. <https://doi.org/10.18653/v1/D19-1018>
 - [29] Ingrid Nunes and Dietmar Jannach. 2017. A Systematic Review and Taxonomy of Explanations in Decision Support and Recommender Systems. *User Modeling and User-Adapted Interaction* 27, 3–5 (Dec. 2017), 393–444. <https://doi.org/10.1007/s11257-017-9195-0>
 - [30] Denis Parra and Peter Brusilovsky. 2015. User-controllable personalization: a case study with SetFusion. *International Journal of Human-Computer Studies* 78 (2015), 43 – 67. <https://doi.org/10.1016/j.ijhcs.2015.01.007>
 - [31] Pearl Pu and Li Chen. 2007. Trust-inspiring explanation interfaces for recommender systems. *Knowledge-Based Systems* 20, 6 (2007), 542 – 556. <https://doi.org/10.1016/j.knsys.2007.04.004>
 - [32] Pearl Pu, Li Chen, and Rong Hu. 2011. A user-centric evaluation framework for recommender systems. In *Proceedings of the Fifth ACM Conference on Recommender Systems* (Chicago, Illinois, USA) (RecSys '11). Association for Computing Machinery, New York, NY, USA, 157–164. <https://doi.org/10.1145/2043932.2043962>
 - [33] Lianping Ren, Hanqin Qiu, Peilai Wang, and Pearl M.C. Lin. 2016. Exploring customer experience with budget hotels: Dimensionality and satisfaction. *International Journal of Hospitality Management* 52 (2016), 13 – 23. <https://doi.org/10.1016/j.ijhm.2015.09.009>
 - [34] Adam Richardson. 2015. Using Customer Journey Maps to Improve Customer Experience. Harvard Business Review.
 - [35] Marc Stickdorn, Jakob Schneider, and Kate Andrews. 2011. *This is service design thinking: Basics, tools, cases*. Wiley.
 - [36] Nava Tintarev and Judith Masthoff. 2012. Evaluating the effectiveness of explanations for recommender systems. *User Modeling and User-Adapted Interaction* 22, 4-5 (2012), 399–439.
 - [37] Nava Tintarev and Judith Masthoff. 2015. *Explaining recommendations: design and evaluation*. Springer US, Boston, MA, 353–382. https://doi.org/10.1007/978-1-4899-7637-6_10
 - [38] Chun-Hua Tsai and Peter Brusilovsky. 2019. Exploring social recommendations with visual diversity-promoting interfaces. *ACM Transactions on Interactive Intelligent Systems* 10, 1, Article 5, 34 pages. <https://doi.org/10.1145/3231465>
 - [39] Katrien Verbert, Denis Parra, and Peter Brusilovsky. 2016. Agents vs. users: visual recommendation of research talks with multiple dimension of relevance. *ACM Transactions on Interactive Intelligent Systems* 6, 2, Article 11 (July 2016), 42 pages. <https://doi.org/10.1145/2946794>
 - [40] Hongwei Wang, Fuzheng Zhang, Jialin Wang, Miao Zhao, Wenjie Li, Xing Xie, and Minyi Guo. 2018. RippleNet: propagating user preferences on the knowledge graph for recommender systems. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management* (Torino, Italy) (CIKM '18). Association for Computing Machinery, New York, NY, USA, 417–426. <https://doi.org/10.1145/3269206.3271739>
 - [41] Wenting Xiong and Diane Litman. 2014. Empirical analysis of exploiting review helpfulness for extractive summarization of online reviews. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*. Dublin City University and Association for Computational Linguistics, Dublin, Ireland, 1985–1995. <https://www.aclweb.org/anthology/C14-1187>
 - [42] Xun Xu and Yibai Lu. 2016. The antecedents of customer satisfaction and dissatisfaction toward various types of hotels: A text mining approach. *International Journal of Hospitality Management* 55 (2016), 57 – 69. <https://doi.org/10.1016/j.ijhm.2016.03.003>