**Beyond Traditional Cultural Heritage Recommender Systems: Suggesting Airbnb Experiences to Users**



(Article begins on next page)

# Beyond Traditional Cultural Heritage Recommender Systems: Suggesting Airbnb Experiences to Users


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

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

Giovanna Petrone
University of Torino
Torino, Italy
giovanna.petrone@unito.it

Angelo Geninatti Cossatin
University of Torino
Torino, Italy
angelo.geninattic@edu.unito.it

Claudio Mattutino
University of Torino
Torino, Italy
claudio.mattutino@unito.it



## ABSTRACT

Traditional recommender systems suggest Cultural and Natural Heritage items to visitors by matching the target user to the available options, one-to-one. However, the increasing diffusion of informal activities and events, supported by location-based services such as Airbnb, extends personalized recommendation to a many-to-one match-making task. Airbnb experiences, which any citizen can propose to offer geographic tours and thematic activities, are composed of at least two entities to be evaluated: the former is the experience itself (in terms of topic, cost, etc.); the latter is the host, who directly interacts with guests during the management of the planned activities. As both entities can dramatically influence guests' perceptions, they should be jointly taken into account by recommender systems. This paper presents our preliminary work aimed at extending the personalized suggestion of Cultural Heritage items to such composite objects.



## CCS CONCEPTS

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

## KEYWORDS

review-based recommender systems, experience modeling, Cultural Heritage

**ACM Reference Format:**
Noemi Mauro, Liliana Ardissono, Giovanna Petrone, Angelo Geninatti Cossatin, and Claudio Mattutino. 2021. Beyond Traditional Cultural Heritage Recommender Systems: Suggesting Airbnb Experiences to Users. 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, 5 pages. https://doi.org/10.1145/3450614.3463385





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





ACM ISBN 978-1-4503-8367-7/21/06…$15.00
https://doi.org/10.1145/3450614.3463385


## 1 INTRODUCTION

The promotion of the Cultural and Natural Heritage of a territory strongly depends on making visitors aware of the offered attractions and events. However, the presence of large sets of options to choose from can overload people with too much data. Several personalized and/or context-aware mobile guides have thus been proposed to help the user focus on the items that (s)he will likely appreciate, by employing recommender systems [33] for information filtering purposes. Typically, these guides suggest physical places, such as museums [3, 6, 20], and public events [7], which they model as a set of features (or tags) that the user might like or dislike. In other words, the item to be suggested is the set of its own properties, which are evaluated to predict the user's interest in it, in a one-to-one recommendation task.

While these models are at the core of the delivery of personalized Cultural Heritage information in traditional settings [4], we claim that they should be extended to exploit the potential of social networks regarding Cultural Heritage promotion. Indeed, the increasing expansion of location-based services, such as Airbnb (https://www.airbnb.com), has made them a novel, low-cost data sharing channel to publish events, activities and services. For instance, Airbnb enables its users to propose *experiences* which people can book, both in their home town and when traveling. Experiences are very heterogeneous and, in our opinion, they are an interesting example of how a territory can be promoted by proposing local activities, products and places in an informal context. For instance, an experience could be a boat tour in a river, the visit to the old town of a city, but also a half day spent with a person who teaches her/his guests how to prepare a traditional dish. It can be noticed that experiences involve a direct interaction between host and guests, as they are directly managed by the organizers.

One-to-one recommendation, in which an item is suggested to a user in a personalized way, is suitable for physical places and structured events. However, we point out that, especially when dealing with informal activities that involve very few people, such as Airbnb experiences, options should be evaluated to estimate both whether the guest will like the topic of the event, and if the interaction with the host is expected to be a good one. In other terms, each experience is more than what is planned for it, and it includes multiple evaluation viewpoints. For instance, the type of activity proposed in an experience might be very interesting for the user but, if the host has been criticized by previous guests, the experience might not represent a good fit. This means that recommendation should be organized as a many-to-one suggestion task, in which the user is matched to a composite object including both the planned activity and the involved actors.

In this work, we explore many-to-one item recommendation by exploring the personalized suggestion of Airbnb experiences. We use the online reviews posted by people on the Airbnb platform in order to extract information about both the experience and its host as separate entities to be evaluated. Moreover, we use the host's description of the experience in the Airbnb web site to address cold start, as this type of data is available as soon as an experience is published. The acquisition of multiple viewpoints on items is the first building block towards the development of recommender systems that will take all these perspectives into account.

This paper describes our preliminary work in this direction. The main contribution is the definition of four models: (i) the experience profile extracted from the description provided by the host, (ii) the one derived from the reviews posted by guests, (iii) the host profile acquired by analyzing her/his self-presentation, and (iv) the one extracted from the comments posted by the people who participated in her/his experiences. All these models are built by extracting features and opinions from textual descriptions that we scraped from the Airbnb web site.

Starting from these models, we plan to build recommender systems that employ different viewpoints on items to generate personalized suggestions. Specifically, we are interested in evaluating whether we can increase recommendation performance by treating host and experience as first-class citizens in item evaluation. Moreover, we plan to fuse hosts and reviewers' perspectives to check whether, in that way, we will improve the resilience to cold start while guaranteeing good performance.

The remainder of this paper is organized as follows: Section 2 outlines the related work. Section 3 describes how the host and experience profiles are built. Section 4 skims the types of recommender systems we plan to develop. Section 5 concludes the paper.

## 2 RELATED WORK

### 2.1 Recommender Systems

Personalized travel guides, such as those described in [6, 7, 20], suggest places and/or events by analyzing the user's interests in item categories (e.g., cinemas or parks), as well as her/his preferences for specific item features, such as the types of artwork exposed in a museum. Moreover, some inclusive recommender systems, like PIUMA [8, 23] and INTRIGUE [3], examine sensory or accessibility features of items to guarantee that the user can smoothly experience the suggested places. In all such cases, the only evaluated entity is the item itself, whose features are matched to the preferences of the target user in order to estimate its suitability.

In Airbnb experiences, the host plays an important role in the success of the interaction with guests. Therefore, different from previous work, we model the host as a separate entity that has to be jointly matched to the user receiving the suggestions in order to evaluate both the offered activity and the overall context in which it will be carried out.

Our work is inspired by review-based recommender systems [9, 17] in what concerns the extraction of data about items from online reviews. As discussed by Ghose and Ipeirotis in [16], consumer feedback is precious, especially when dealing with experience goods and services, because it conveys information about how people perceive items. It can thus be employed to complement metadata with first-hand opinions. However, review-based recommender systems support a one-to-one type of matching between users and items. For example, in [10, 12, 22, 29], the user model is instantiated on the basis of the reviews (s)he posts. Moreover, the item model depends on the overall amount of reviews it collects. Differently, our model distinguishes the multiple entities involved in an experience (i.e., host and activity to be performed) and it separately extracts information about all of them to support a many-to-one type of recommendation. Moreover, we consider both the perspective of the host and that of the guest in the representation of items.

### 2.2 Feature Extraction Techniques

In Content-Based Filtering, item features are typically extracted from textual catalogs by applying statistical metrics like TF-IDF. Moreover, in review-based recommender systems, opinion mining techniques such as faceted opinion extraction [26], bi-gram and tri-gram analysis [11], Non-negative Matrix Factorization [5], Latent Dirichlet Allocation (LDA) [2, 25] and *ensemble* methods [29] are used to identify aspects in reviews. Further techniques are applied in the extraction of sentiment about products and services. For instance, Qi et al. [30] combine LDA with PageRank on terms to find relevant product properties, and Korfiatis et al. [19] apply Structural Topic Models to identify evaluation aspects in reviews. Finally, Paul et al., [28] use Double Propagation [31] and Xu et al. [36] apply Latent Semantic Analysis to extract aspects from reviews as latent topics.

We apply Double Propagation because it supports a flexible and extensible extraction of aspects from articulated sentences. However, we separate the aspects concerning the host from those related to the proposed activity in order to support the creation of multiple item and host profiles.

## 3 MODEL

### 3.1 Dataset

During April and May 2019 we scraped the Airbnb web site to collect information about the experiences it offered, the associated reviews and their hosts, focusing on content written in English. For the scraping task, we built a bot that, by exploiting the Selenium Python library [27], analyzed the HTML web pages. Moreover, we used the Langdetect Python library [34] to filter English content. We organized the collected data in two CSV documents which respectively described the experiences and the hosts.
Each experience $E$ is has several fields, among which:

- ID, URL, type of experience, price, offered amenities, "what-we-will-do" field specifying the type of activity that will be carried out, expected duration of $E$, "where-we-will-be" field describing the location of $E$, city and country in which $E$ is offered, and language that will be used during the activity.
- ID, name and URL of the host offering $E$, and the host's description about her/himself.

**Table 1: Aspects of an Airbnb experience, extracted from the filtered Airbnb dataset.**

| Aspect | Adjective | Negation | Score |
|---|---|---|---|
| boat | big | false | 3.00 |
| experience | amazing | false | 4.19 |
| tour | great | false | 4.43 |
| snack | great | false | 4.43 |
| snack | good | false | 4.14 |
| experience | wonderful | false | 4.57 |
| experience | intimate | false | 3.2 |

**Table 2: Aspects about the host of the experience of Table 1.**

| Aspect | Adjective | Negation | Score |
|---|---|---|---|
| he | great | false | 4.43 |
| host | friendly | false | 3.87 |
| guide | amazing | false | 4.19 |
| host | knowledgeable | false | 3.00 |
| Stephan | friendly | false | 3.87 |
| host | funny | false | 3.69 |
| host | lovely | false | 4.08 |

- Overall number of reviews about $E$ in the Airbnb web site, and mean rating in [1, 5] given by guests to $E$.

Each review $R$ is described by the following fields:

- Review ID and date, ID of the experience $E$ to which $R$ refers.
- ID and name of $R$'s author.
- Textual comment of $R$, and score that the reviewer gave to $E$, in [1, 5].

We collected data about 254,253 Airbnb users, 11,086 experiences and 336,288 reviews. However, we noticed that most users wrote very few reviews. We thus filtered the dataset on the users who provided at least 5 reviews. The filtered dataset contains 2,386 users.

It can be noticed that not all the fields of the descriptors are structured. For instance, the date of a review, the price and the type of an experience store tokens; however, other fields, such as the "what-we-will-do" one, contain Natural Language text that has to be analyzed to obtain a structured representation ready to be used for recommendation purposes. Section 3.2 outlines this analysis process.

## 3.2 Extraction of Features about Hosts and Experiences

We extract hosts and experiences' features from the textual descriptions of the filtered dataset by applying standard NLP techniques aimed at recognizing the mentioned aspects and their values. Given the set of sentences to be analyzed, e.g., all the reviews about an individual experience, we group the extracted aspects to distinguish those referring to the experience from those referring to its host. This is key to build separate models for these two types of entities.

For the extraction of aspects from a text, we use the Spacy dependency parser [14], which builds the syntax tree of each sentence, and we lemmatize the words to have a single representation of terms. Starting from the syntax tree, we apply the rules described in the Double Propagation algorithm by Qiu et al. [32] to identify aspects and corresponding adjectives. In order to separate the aspects concerning the host from those about the experience, we use a dictionary of terms that usually refer to the host. These include proper nouns, pronouns such as "he" and "she" and a set of nouns referring to host roles such as "driver", "guide", and so forth. We built this dictionary by collecting the noun phrases referring to hosts in a sample set of Airbnb reviews.

Table 1 shows the aspects of an individual experience $E$ that we extracted from the set of reviews about $E$ stored in our dataset. Similarly, Table 2 shows the aspects concerning the host of $E$. Each row of the tables shows an aspect $A$, an adjective associated to $A$ in the analyzed sentence, a marker that specifies whether the adjective is negated or not, and the resulting evaluation score of $A$. We obtain the score of each aspect as follows: first, we compute the polarity of each adjective as the mean value of the polarities returned by the TextBlob [21] and Vader [18] opinion mining libraries, taking the presence of negations into account. Then, we normalize this value in the [1, 5] interval, which we use for rating prediction.

It can be noticed that the tables contain multiple occurrences of the aspects that are mentioned more than once in the analyzed text. Moreover, in principle, the same aspects might be described using synonyms. For instance, in other experiences we found the occurrence of both "snack" and "food", or "tour" and "trip". Currently, we overlook this issue and we treat each lemmatized term as a different aspect. However, in order to enhance the recognition of multiple occurrences of the same aspects, we plan to normalize data by using synonyms and by merging similar aspects on the basis of their semantic similarity.

## 3.3 Host and Experience Models for Recommendation

Reviewers' feedback is a valuable information source to evaluate experiences because it describes their participants' opinions about them. However, new experiences need some time to receive their first comments. Moreover, reviews typically focus on the main perceptions about items and, as such, they might have to be complemented with other data to build rich item models. In order to address these issues, we propose to extract the aspects of hosts and experiences by analyzing both the online reviews, and the descriptions published by the hosts. For each experience $E$, we build four item profiles that characterize $E$ and its host $H$:

- *Host-by-reviewers:* this profile is acquired by analyzing all the reviews about the experiences offered by $H$, and by extracting the comments that concern her/him. It thus represents the opinions expressed by her/his guests.
- *Experience-by-reviewers:* this profile is acquired by analyzing all the online reviews about $E$, and by extracting the comments that concern it.
- *Host-by-host:* this profile is acquired by analyzing the description that $H$ provides about her/himself.
- *Experience-by-host:* this profile is built by analyzing the description of $E$ provided by $H$.

Each item profile is a vector $<< a_1, meanScore_1 >, \ldots, < a_n, meanScore_n >>$, where, for $j$ in [1, n]:

- $a_j$ is an aspect of the described entity;
- $meanScore_j$ is the mean evaluation score of $a_j$ in the analyzed text.

For instance, the experience $E$ of Table 1 has the following *Experience-by-reviewers* profile: $<< boat, 3.00 >, < experience, 3.987 >, \ldots, < snack, 4.285 >>$. Moreover, the *Host-by-reviewers* profile of $E$'s host (see Table 2) is $<< host, 3.876 >>$. In this case, the aspects are fused into a single pair because they all refer to $E$'s host, regardless of the name or pronoun used to refer to him.

### 3.4 User Profiles

We plan to estimate the ratings of items based on the scores of aspects in their profiles, and of the importance that the user gives to such aspects. Our idea is that, if a host **h** gets good scores in aspects that a user **u** considers as important, then it is likely that **u** will appreciate **h** in the experiences (s)he hosts. Similarly, the importance of the aspects of experiences can be employed to evaluate their suitability for the user. We thus define the user profile as a vector of pairs: $<< a_1, importance_1 >, \ldots, < a_n, importance_n >>$, where the aspects $a_j$ concern either hosts or experiences, and the importance values are decimal numbers in [0, 1].

## 4 HOST AND EXPERIENCE-AWARE RECOMMENDER SYSTEMS

### 4.1 One-to-one Recommendation

A simple way to personalized the suggestion of experiences to a user **u** is the adoption of an evaluation model based on her/his user profile and on one of the types of item profile described in Section 3.3; i.e., either *Experience-by-reviewers* or *Experience-by-host*. Given a user profile **u**, and an item profile **i**, we can estimate the rating of **i** as the weighted mean of the scores of its aspects, tuned by the importance that **u** gives to such aspects:

$$\frac{\sum_{j=1}^{n} importance_{uj} * score_{ij}}{n} \tag{1}$$

where $importance_{uj}$ is the importance of aspect $a_j$ in the user profile and $score_{ij}$ is the evaluation score of $a_j$ in the item profile. This is a one-to-one match-making, which does not support the evaluation of composite objects such as experiences. However, it is the basis for the integration of multiple perspectives for item recommendation.

### 4.2 Many-to-one Recommendation

The separation of host and experience-related aspects supports the design of recommendation strategies that are partially related to multi-criteria recommender systems [1]. The idea is that, as both the host and the activity to be carried out influence the guest's perception of the overall experience, they represent two main entities to be matched against the preferences of the target user. In this respect, we are interested in developing recommendation algorithms by considering different design dimensions for the evaluation of their performance:

- **Type of match-making: one-to-one vs. many-to-one.** We plan to compare the recommendation performance of algorithms focusing on the model of the experience with that of algorithms that also take the host model into account. For instance, in the latter case, the recommender system might separately evaluate the experience and the host. Then, it might merge these evaluations, e.g., by computing their weighted mean, or their Fuzzy AND, depending on the required selectivity of the suggestions.
- **Viewpoint in the descriptions of hosts and experiences.** As previously mentioned, for each host and experience there are two models: one built from the reviews, the other one derived from the descriptions provided by the host. We are interested in evaluating the performance of a recommender system that only relies on reviewers' feedback, with respect to another one that merges this feedback with the hosts' descriptions, with specific attention to cold start.

We aim at carrying out a user test to investigate user experience with these recommender systems. However, a first type of validation can be an offline one that leverages the data stored in the filtered Airbnb dataset to build item and user profiles, and to measure recommendation performance using the observed item ratings. We can build the user profiles of each user **u** by analyzing the reviews provided by her/him in order to infer the importance (s)he gives to the various aspects of items and hosts. For this purpose, we can merge all the reviews posted by **u** and analyze their text, using the same techniques described in Section 3.2. The result of this analysis are two tables reporting the mentioned aspects and the number of occurrences of the aspects respectively concerning experiences and hosts. By normalizing the number of occurrences of items in [0, 1], we can derive the relative importance of aspects to **u**, and thus her/his user profile.

## 5 CONCLUSIONS

The promotion of Cultural Heritage is becoming more inclusive thanks to the presence of location-based services that empower individual citizens to propose informal activities rooted in the territory. For instance, thanks to the Airbnb experiences, people can offer tours of a geographic area, as well as other types of events, which they carry out with their guests. However, this fluid context changes the setting of recommender systems. Specifically, when suggesting an experience to a person, the system should take care that (s)he appreciates both the topic of the experience and its host, as both of them might contribute to spoil it, or to make it a very good one. In other words, the proposed item is composed of at least two entities to be evaluated. In this paper, we described our preliminary work towards the development of recommender systems that manage a many-to-one type of match-making between users and items, by modeling items as composite sets of entities. We outlined the models that come into play in this context and the type of algorithms that might be designed.

Our future work includes the development of the envisaged algorithms and their evaluation, offline and with real users. Moreover, we plan to enhance the extraction of aspects from reviews by using review helpfulness analysis [13, 15, 24, 35] to select valuable consumer feedback and filter out low-quality reviews.

## ACKNOWLEDGMENTS

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

## REFERENCES


[1] Gediminas Adomavicius, Nikos Manouselis, and YoungOk Kwon. 2011. *Multi-Criteria Recommender Systems*. Springer US, Boston, MA, 769–803. https://doi.org/10.1007/978-0-387-85820-3_24

[2] Marie Al-Ghossein, Pierre-Alexandre Murena, Talel Abdessalem, Anthony Barré, and Antoine Cornuéjols. 2018. Adaptive Collaborative Topic Modeling for Online Recommendation. In *Proceedings of the 12th ACM Conference on Recommender Systems* (Vancouver, British Columbia, Canada) *(RecSys '18)*. Association for Computing Machinery, New York, NY, USA, 338–346. https://doi.org/10.1145/3240323.3240363

[3] Liliana Ardissono, Anna Goy, Giovanna Petrone, Marino Segnan, and Pietro Torasso. 2003. INTRIGUE: personalized recommendation of tourist attractions for desktop and handset devices. *Applied Artificial Intelligence, Special Issue on Artificial Intelligence for Cultural Heritage and Digital Libraries* 17, 8-9 (2003), 687–714. https://doi.org/10.1080/713827254

[4] L. Ardissono, D. Petrelli, and T. Kuflik. 2012. Personalization in Cultural Heritage: The Road Travelled and the One Ahead. *USER MODELING AND USER-ADAPTED INTERACTION - The Journal of Personalization Research. Special Issue on Coming of Age: Celebrating a Quarter Century of User Modeling and Personalization* 22, 1-2 (2012), 73–99.

[5] Yang Bao, Hui Fang, and Jie Zhang. 2014. TopicMF: simultaneously exploiting ratings and reviews for recommendation. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence* (Québec City, Québec, Canada) *(AAAI'14)*. AAAI Press, 2–8. https://doi.org/10.5555/2893873.2893874

[6] Matthias Braunhofer and Francesco Ricci. 2016. Contextual Information Elicitation in Travel Recommender Systems. In *Information and Communication Technologies in Tourism 2016*, Alessandro Inversini and Roland Schegg (Eds.). Springer International Publishing, Cham, 579–592. https://doi.org/10.1007/978-3-319-28231-2_42

[7] F. Carmagnola, F. Cena, L. Console, O. Cortassa, C. Gena, A. Goy, I. Torre, A. Toso, and F. Vernero. 2008. Tag-based user modeling for social multi-device adaptive guides. *User Modeling and User-Adapted Interaction* 18, 5 (2008), 497–538.

[8] Federica Cena, Noemi Mauro, Liliana Ardissono, Claudio Mattutino, Amon Rapp, Stefano Cocomazzi, Stefania Brighenti, and Roberto Keller. 2020. Personalized Tourist Guide for People with Autism. In *Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (Genoa, Italy) *(UMAP '20 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 347–351. https://doi.org/10.1145/3386392.3399280

[9] Li Chen, Guanliang Chen, and Feng Wang. 2015. Recommender systems based on user reviews: the state of the art. *User Modeling and User-Adapted Interaction* 25, 2 (2015), 99–154. https://doi.org/10.1007/s11257-015-9155-5

[10] Ruihai Dong, Michael P. O'Mahony, Markus Schaal, Kevin Mccarthy, and Barry Smyth. 2016. Combining Similarity and Sentiment in Opinion Mining for Product Recommendation. *Journal of Intelligent Information Systems* 46, 2 (April 2016), 285–312. https://doi.org/10.1007/s10844-015-0379-y

[11] Ruihai Dong, Markus Schaal, Michael P. O'Mahony, and Barry Smyth. 2013. Topic Extraction from Online Reviews for Classification and Recommendation. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence* (Beijing, China) *(IJCAI '13)*. AAAI Press, 1310–1316.

[12] Ruihai Dong and Barry Smyth. 2017. User-based Opinion-based Recommendation. In *Proceedings 26th IJCAI*. Melbourne, Australia, 4821–4825.

[13] Seyed Pouyan Eslami, Maryam Ghasemaghaei, and Khaled Hassanein. 2018. Which online reviews do consumers find most helpful? A multi-method investigation. *Decision Support Systems* 113 (2018), 32 – 42. https://doi.org/10.1016/j.dss.2018.06.012

[14] Explosion AI. 2017. SpaCy - Industrial Natural Language Processing in Python. https://spacy.io/.

[15] Jorge E. Fresneda and David Gefen. 2019. A semantic measure of online review helpfulness and the importance of message entropy. *Decision Support Systems* 125 (2019), 113117. https://doi.org/10.1016/j.dss.2019.113117

[16] Anindya Ghose and Panagiotis G. Ipeirotis. 2011. Estimating the helpfulness and economic impact of product reviews: mining text and reviewer characteristics. *IEEE Transactions on on Knowledge and Data Engineering* 23, 10 (2011), 1498–1512. https://doi.org/10.1109/TKDE.2010.188

[17] María Hernández-Rubio, Iván Cantador, and Alejandro Bellogín. 2019. A comparative analysis of recommender systems based on item aspect opinions extracted from user reviews. *User Modeling and User-Adapted Interaction* 29, 2 (2019), 381–441. https://doi.org/10.1007/s11257-018-9214-9

[18] C.J. Hutto and Gilbert Eric. 2014. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. In *Proceedings of the 8th International AAAI Conference on Weblogs and Social Media* (Beijing, China). AAAI, New York, NY, USA, 216–225. https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/viewPaper/8109

[19] Nikolaos Korfiatis, Panagiotis Stamolampros, Panos Kourouthanassis, and Vasileios Sagiadinos. 2019. Measuring service quality from unstructured data: A topic modeling application on airline passengers' online reviews. *Expert Systems with Applications* 116 (2019), 472 – 486. https://doi.org/10.1016/j.eswa.2018.09.037

[20] A. Krüger, J. Baus, D. Heckmann, M. Kruppa, and R. Wasinger. 2007. Web-based mobile guides. In *The Adaptive Web: Methods and Strategies of Web Personalization, Lecture Notes in Computer Science, Vol. 4321*, P. Brusilovsky, A. Kobsa, and W. Nejdl (Eds.). Springer-Verlag, Berlin Heidelberg New York, 521–549.

[21] Steven Loria. 2020. TextBlob: Simplified Text Processing. https://textblob.readthedocs.io/en/dev/index.html.

[22] Yichao Lu, Ruihai Dong, and Barry Smyth. 2018. Coevolutionary Recommendation Model: Mutual Learning between Ratings and Reviews. In *Proceedings of the 2018 World Wide Web Conference* (Lyon, France) *(WWW '18)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 773–782. https://doi.org/10.1145/3178876.3186158

[23] Noemi Mauro, Liliana Ardissono, and Federica Cena. 2020. Personalized Recommendation of PoIs to People with Autism. In *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (Genova, Italy) *(UMAP'20)*. ACM, New York, NY, USA, 163–172. https://doi.org/10.1145/3340631.3394845

[24] Noemi Mauro, Liliana Ardissono, and Giovanna Petrone. 2021. User and item-aware estimation of review helpfulness. *Information Processing & Management* 58, 1 (2021), 102434. https://doi.org/10.1016/j.ipm.2020.102434

[25] Julian McAuley and Jure Leskovec. 2013. Hidden factors and hidden topics: understanding rating dimensions with review text. In *Proceedings of the 7th ACM Conference on Recommender Systems* (Hong Kong, China) *(RecSys '13)*. Association for Computing Machinery, New York, NY, USA, 165–172. https://doi.org/10.1145/2507157.2507163

[26] Claudiu-Cristian Musat and Boi Faltings. 2015. Personalizing product rankings using collaborative filtering on opinion-derived topic profiles. In *Proceedings of the 24th International Conference on Artificial Intelligence* (Buenos Aires, Argentina) *(IJCAI'15)*. AAAI Press, 830–836. https://dl.acm.org/doi/10.5555/2832249.2832364

[27] Baiju Muthukadan. 2019. Selenium with Python. https://selenium-python.readthedocs.io/.

[28] Debanjan Paul, Sudeshna Sarkar, Muthusamy Chelliah, Chetan Kalyan, and Prajit Prashant Sinai Nadkarni. 2017. Recommendation of High Quality Representative Reviews in E-Commerce. In *Proceedings of the Eleventh ACM Conference on Recommender Systems* (Como, Italy) *(RecSys '17)*. Association for Computing Machinery, New York, NY, USA, 311–315. https://doi.org/10.1145/3109859.3109901

[29] Francisco J. Peña, Diarmuid O'Reilly-Morgan, Elias Z. Tragos, Neil Hurley, Erika Duriakova, Barry Smyth, and Aonghus Lawlor. 2020. Combining Rating and Review Data by Initializing Latent Factor Models with Topic Models for Top-N Recommendation. In *Fourteenth ACM Conference on Recommender Systems* (Virtual Event, Brazil) *(RecSys '20)*. Association for Computing Machinery, New York, NY, USA, 438–443. https://doi.org/10.1145/3383313.3412207

[30] Jiayin Qi, Zhenping Zhang, Seongmin Jeon, and Yanquan Zhou. 2016. Mining customer requirements from online reviews: A product improvement perspective. *Information & Management* 53, 8 (2016), 951 – 963. https://doi.org/10.1016/j.im.2016.06.002

[31] Guang Qiu, Bing Liu, Jiajun Bu, and C. Chen. 2011. Opinion Word Expansion and Target Extraction through Double Propagation. *Computational Linguistics* 37 (03 2011), 9–27. https://doi.org/10.1162/coli_a_00034

[32] L. Qiu, L. Chang, F. Lin, and Z. Shi. 2007. Context Optimization of AI Planning for Semantic Web Services Composition. *Service Oriented Computing and Applications* 1, 2 (2007), 117–128.

[33] Francesco Ricci, Lior Rokach, and Bracha Shapira. 2011. *Introduction to Recommender Systems Handbook*. Springer US, Boston, MA, 1–35. https://doi.org/10.1007/978-0-387-85820-3_1

[34] Nakatani Shuyo. 2020. langdetect. https://pypi.org/project/langdetect/.

[35] Xinyu Sun, Maoxin Han, and Juan Feng. 2019. Helpfulness of online reviews: examining review informativeness and classification thresholds by search products and experience products. *Decision Support Systems* 124 (2019), 113099. https://doi.org/10.1016/j.dss.2019.113099

[36] Xun Xu, Xuequn Wang, Yibai Li, and Mohammad Haghighi. 2017. Business intelligence in online customer textual reviews: understanding consumer perceptions and influential factors. *International Journal of Information Management* 37, 6 (2017), 673 – 683. https://doi.org/10.1016/j.ijinfomgt.2017.06.004