Explainable Recommendation: A Survey and New Perspectives

Other titles in Foundations and Trends® in Information Retrieval

Information Retrieval: The Early Years
Donna Harman
ISBN: 978-1-68083-584-7

Bandit Algorithms in Information Retrieval
Dorota Glowacka
ISBN: 978-1-68083-574-8

Neural Approaches to Conversational AI
Jianfeng Gao, Michel Galley and Lihong Li
ISBN: 978-1-68083-552-6

An Introduction to Neural Information Retrieval
Bhaskar Mitra and Nick Craswell
ISBN: 978-1-68083-532-8

Explainable Recommendation: A Survey and New Perspectives

Yongfeng Zhang
Rutgers University, USA
yongfeng.zhang@rutgers.edu

Xu Chen
Tsinghua University, China
xu-ch14@mails.tsinghua.edu.cn

the essence of knowledge
Boston — Delft

Foundations and Trends® in Information Retrieval

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

Y. Zhang and X. Chen. *Explainable Recommendation: A Survey and New Perspectives*. Foundations and Trends® in Information Retrieval, vol. 14, no. 1, pp. 1–101, 2020.

ISBN: 978-1-68083-659-2

Foundations and Trends® in Information Retrieval
Volume 14, Issue 1, 2020

Editorial Scope

Topics

Foundations and Trends® in Information Retrieval publishes survey and tutorial articles in the following topics:

- Applications of IR
- Architectures for IR
- Collaborative filtering and recommender systems
- Cross-lingual and multilingual IR
- Distributed IR and federated search
- Evaluation issues and test collections for IR
- Formal models and language models for IR
- IR on mobile platforms
- Indexing and retrieval of structured documents
- Information categorization and clustering
- Information extraction
- Information filtering and routing
- Metasearch, rank aggregation and data fusion
- Natural language processing for IR
- Performance issues for IR systems, including algorithms, data structures, optimization techniques, and scalability
- Question answering
- Summarization of single documents, multiple documents, and corpora
- Text mining
- Topic detection and tracking
- Usability, interactivity, and visualization issues in IR
- User modelling and user studies for IR
- Web search

Information for Librarians

Foundations and Trends® in Information Retrieval, 2020, Volume 14, 5 issues. ISSN paper version 1554-0669. ISSN online version 1554-0677. Also available as a combined paper and online subscription.

Contents

1 Introduction **3**
1.1 Explainable Recommendation 3
1.2 A Historical Overview 4
1.3 Classification of the Methods 7
1.4 Explainability and Effectiveness 10
1.5 Explainability and Interpretability 10
1.6 How to Read the Survey 11

2 Information Source for Explanations **13**
2.1 Relevant User or Item Explanation 14
2.2 Feature-based Explanation 17
2.3 Opinion-based Explanation 19
2.4 Sentence Explanation 21
2.5 Visual Explanation 24
2.6 Social Explanation 26
2.7 Summary 28

3 Explainable Recommendation Models **29**
3.1 Overview of Machine Learning for Recommendation 30
3.2 Factorization Models for Explainable Recommendation . . 31
3.3 Topic Modeling for Explainable Recommendation 35
3.4 Graph-based Models for Explainable Recommendation . . . 38

3.5 Deep Learning for Explainable Recommendation 40
3.6 Knowledge Graph-based Explainable Recommendation . . . 46
3.7 Rule Mining for Explainable Recommendation 49
3.8 Model Agnostic and Post Hoc Explainable Recommendation 51
3.9 Summary 54

4 Evaluation of Explainable Recommendation 56
4.1 User Study 57
4.2 Online Evaluation 59
4.3 Offline Evaluation 61
4.4 Qualitative Evaluation by Case Study 62
4.5 Summary 64

5 Explainable Recommendation in Different Applications 65
5.1 Explainable E-commerce Recommendation 65
5.2 Explainable Point-of-Interest Recommendation 66
5.3 Explainable Social Recommendation 68
5.4 Explainable Multimedia Recommendation 69
5.5 Other Explainable Recommendation Applications 70
5.6 Summary 71

6 Open Directions and New Perspectives 72
6.1 Methods and New Applications 72
6.2 Evaluation and User Behavior Analysis 77
6.3 Explanation for Broader Impacts 78
6.4 Cognitive Science Foundations 79

7 Conclusions 80

Acknowledgements 82

References 83

Explainable Recommendation: A Survey and New Perspectives

Yongfeng Zhang[1] and Xu Chen[2]

[1] *Rutgers University, USA; yongfeng.zhang@rutgers.edu*
[2] *Tsinghua University, China; xu-ch14@mails.tsinghua.edu.cn*

ABSTRACT

Explainable recommendation attempts to develop models that generate not only high-quality recommendations but also intuitive explanations. The explanations may either be post-hoc or directly come from an explainable model (also called interpretable or transparent model in some contexts). Explainable recommendation tries to address the problem of *why*: by providing explanations to users or system designers, it helps humans to understand why certain items are recommended by the algorithm, where the human can either be users or system designers. Explainable recommendation helps to improve the transparency, persuasiveness, effectiveness, trustworthiness, and satisfaction of recommendation systems. It also facilitates system designers for better system debugging. In recent years, a large number of explainable recommendation approaches – especially model-based methods – have been proposed and applied in real-world systems.

In this survey, we provide a comprehensive review for the explainable recommendation research. We first highlight the position of explainable recommendation in recommender system research by categorizing recommendation problems into the 5W, i.e., what, when, who, where, and why. We then

Yongfeng Zhang and Xu Chen (2020), “Explainable Recommendation: A Survey and New Perspectives”, Foundations and Trends® in Information Retrieval: Vol. 14, No. 1, pp 1–101. DOI: 10.1561/1500000066.

conduct a comprehensive survey of explainable recommendation on three perspectives: 1) We provide a chronological research timeline of explainable recommendation, including user study approaches in the early years and more recent model-based approaches. 2) We provide a two-dimensional taxonomy to classify existing explainable recommendation research: one dimension is the information source (or display style) of the explanations, and the other dimension is the algorithmic mechanism to generate explainable recommendations. 3) We summarize how explainable recommendation applies to different recommendation tasks, such as product recommendation, social recommendation, and POI recommendation.

We also devote a section to discuss the explanation perspectives in broader IR and AI/ML research. We end the survey by discussing potential future directions to promote the explainable recommendation research area and beyond.

1

Introduction

1.1 Explainable Recommendation

Explainable recommendation refers to personalized recommendation algorithms that address the problem of *why* – they not only provide users or system designers with recommendation results, but also explanations to clarify why such items are recommended. In this way, it helps to improve the transparency, persuasiveness, effectiveness, trustworthiness, and user satisfaction of the recommendation systems. It also facilitates system designers to diagnose, debug, and refine the recommendation algorithm.

To highlight the position of explainable recommendation in the recommender system research area, we classify personalized recommendation with a broad conceptual taxonomy. Specifically, personalized recommendation research can be classified into the 5W problems – when, where, who, what, and why, corresponding to time-aware recommendation (when), location-based recommendation (where), social recommendation (who), application-aware recommendation (what), and explainable recommendation (why), where explainable recommendation aims to answer *why*-type questions in recommender systems.

Explainable recommendation models can either be model-intrinsic or model-agnostic (Lipton, 2018; Molnar, 2019). The model-intrinsic approach develops interpretable models, whose decision mechanism is transparent, and thus, we can naturally provide explanations for the model decisions (Zhang *et al.*, 2014a). The model-agnostic approach (Wang *et al.*, 2018d), or sometimes called the post-hoc explanation approach (Peake and Wang, 2018), allows the decision mechanism to be a blackbox. Instead, it develops an explanation model to generate explanations after a decision has been made. The philosophy of these two approaches is deeply rooted in our understanding of human cognitive psychology – sometimes we make decisions by careful, rational reasoning and we can explain why we make certain decisions; other times we make decisions first and then find explanations for the decisions to support or justify ourselves (Lipton, 2018; Miller, 2019).

The scope of explainable recommendation not only includes developing transparent machine learning, information retrieval, or data mining models. It also includes developing effective methods to deliver the recommendations or explanations to users or system designers, because explainable recommendations naturally involve humans in the loop. Significant research efforts in user behavior analysis and human-computer interaction community aim to understand how users interact with explanations.

With this section, we will introduce not only the explainable recommendation problem, but also a big picture of the recommender system research area. It will help readers to understand what is unique about the explainable recommendation problem, what is the position of explainable recommendation in the research area, and why explainable recommendation is important to the area.

1.2 A Historical Overview

In this section, we will provide a historical overview of the explainable recommendation research. Though the term *explainable recommendation* was formally introduced in recent years (Zhang *et al.*, 2014a), the basic concept, however, dates back to some of the earliest works in personalized recommendation research. For example, Schafer *et al.* (1999) noted that recommendations could be explained by other items that the user is familiar with, such as *this product you are looking at is similar to these other products you liked before*, which leads to the fundamental idea of item-based collaborative filtering (CF); Herlocker *et al.* (2000) studied how to explain CF algorithms in MovieLens based on user surveys; and Sinha and Swearingen (2002) highlighted the role of transparency in recommender systems. Besides, even before explainable recommendation has attracted serious research attention, the industry has been using manual or semi-automatic explanations in practical systems, such as the *people also viewed* explanation in e-commerce systems (Tintarev and Masthoff, 2007a).

To help the readers understand the "pre-history" research of recommendation explanation and how explainable recommendation emerged as an essential research task in the recent years, we provide a historical overview of the research line in this section.

Early approaches to personalized recommender systems mostly focused on content-based or collaborative filtering (CF)-based recommendation (Ricci *et al.*, 2011). Content-based recommender systems model user and item profiles with various available content information, such as the price, color, brand of the goods in e-commerce, or the genre, director, duration of the movies in review systems (Balabanović and Shoham, 1997; Pazzani and Billsus, 2007). Because the item contents are easily understandable to users, it was usually intuitive to explain to users why an item is recommended. For example, one straightforward way is to let users know the content features he/she might be interested in the recommended item. Ferwerda *et al.* (2012) provided a comprehensive study of possible protocols to provide explanations for content-based recommendations.

However, collecting content information in different application domains is time-consuming. Collaborative filtering (CF)-based approaches (Ekstrand *et al.*, 2011), on the other hand, attempt to avoid this difficulty by leveraging *the wisdom of crowds*. One of the earliest CF algorithms is User-based CF for the GroupLens news recommendation system (Resnick *et al.*, 1994). User-based CF represents each user as a vector of ratings, and predicts the user's missing rating on a news message based on the weighted average of other users' ratings on the message. Symmetrically, Sarwar *et al.* (2001) introduced the Item-based CF method, and Linden *et al.* (2003) further described its application in Amazon product recommendation system. Item-based CF takes each item as a vector of ratings, and predicts the missing rating based on the weighted average of ratings from similar items.

Though the rating prediction mechanism would be relatively difficult to understand for average users, user- and item-based CF are somewhat explainable due to the philosophy of their algorithm design. For example, the items recommended by user-based CF can be explained as "users that are similar to you loved this item", while item-based CF can be explained as "the item is similar to your previously loved items". However, although the idea of CF has achieved significant improvement in recommendation accuracy, it is less intuitive to explain compared with content-based algorithms. Research pioneers in very early stages also noticed the importance of the problem (Herlocker and Konstan, 2000; Herlocker *et al.*, 2000; Sinha and Swearingen, 2002).

The idea of CF achieved further success when integrated with Latent Factor Models (LFM) introduced by Koren (2008) in the late 2000s. Among the many LFMs, Matrix Factorization (MF) and its variants were especially successful in rating prediction tasks (Koren *et al.*, 2009). Latent factor models have been leading the research and application of recommender systems for many years. However, though successful in recommendation performance, the "latent factors" in LFMs do not possess intuitive meanings, which makes it difficult to understand why an item got good predictions or why it got recommended out of other candidates. This lack of model explainability also makes it challenging to provide intuitive explanations to users, since it is hardly acceptable to tell users that we recommend an item only because it gets higher prediction scores by the model.

To make recommendation models better understandable, researchers have gradually turned to *Explainable Recommendation Systems*, where the recommendation algorithm not only outputs a recommendation list, but also explanations for the recommendations by working in an explainable way. For example, Zhang *et al.* (2014a) defined the *explainable recommendation* problem, and proposed an Explicit Factor

Model (EFM) by aligning the latent dimensions with explicit features for explainable recommendation. More approaches were also proposed to address the explainability problem, which we will introduce in detail in the survey. It is worthwhile noting that deep learning (DL) models for personalized recommendation have emerged in recent years. We acknowledge that whether DL models truly improve the recommendation performance is controversial (Dacrema *et al.*, 2019), but this problem is out of the scope of this survey. In this survey, we will focus on the problem that the black-box nature of deep models brings difficulty in model explainability. We will review the research efforts on explainable recommendation over deep models.

In a broader sense, the explainability of AI systems was already a core discussion in the 1980s era of "old" or logical AI research, when knowledge-based systems predicted (or diagnosed) well but could not explain why. For example, the work of Clancy showed that being able to explain predictions requires far more knowledge than just making correct predictions (Clancey, 1982). The recent boom in big data and computational power have brought AI performance to a new level, but researchers in the broader AI community have again realized the importance of *Explainable AI* in recent years (Gunning, 2017), which aims to address a wide range of AI explainability problems in deep learning, computer vision, autonomous driving systems, and natural language processing tasks. As an essential branch of AI research, this also highlights the importance of the IR/RecSys community to address the explainability issues of various search and recommendation systems. Moreover, explainable recommendation has also become a very suitable problem setting to develop new *Explainable Machine Learning* theories and algorithms.

1.3 Classification of the Methods

In this survey, we provide a classification taxonomy of existing explainable recommendation methods, which can help readers to understand the state-of-the-art of explainable recommendation research.

Specifically, we classify existing explainable recommendation research with two orthogonal dimensions: 1) The information source or display style of the explanations (e.g., textual sentence explanation, or visual explanation), which represents the human-computer interaction (HCI) perspective of explainable recommendation research, and 2) the model to generate such explanations, which represents the machine learning (ML) perspective of explainable recommendation research. Potential explainable models include the nearest-neighbor, matrix factorization, topic modeling, graph models, deep learning, knowledge reasoning, association rule mining, and others.

With this taxonomy, each combination of the two dimensions refers to a particular sub-direction of explainable recommendation research. We should note that there could exist conceptual differences between "how explanations are presented (display style)" and "the type of information used for explanations (information source)". In the context of explainable recommendation, however, these two principles are closely related to each other because the type of information usually determines how the explanations can be displayed. As a result, we merge these two principles into a single classification dimension. Note that among the possibly many classification taxonomies, this is just one that we think would be appropriate to organize the research on explainable recommendation, because it considers both HCI and ML perspectives of explainable recommendation research.

Table 1.1 shows how representative explainable recommendation research is classified into different categories. For example, the Explicit Factor Model (EFM) for explainable recommendation (Zhang *et al.*, 2014a) developed a matrix factorization method for explainable recommendation, which provides an explanation sentence for the recommended item. As a result, it falls into the category of "matrix factorization with textual explanation". The Interpretable Convolutional Neural Network approach (Seo *et al.*, 2017), on the other hand, develops a deep convolutional neural network model and displays item features to users as explanations, which falls into the category of "deep learning with user/item feature explanation". Another example is visually explainable recommendation (Chen *et al.*, 2019b), which proposes a deep model to generate image regional-of-interest explanations, and it belongs to the "deep learning with visual explanation" category. We also classify other

Table 1.1: A classification of existing explainable recommendation methods. The classification is based on two dimensions, i.e., the type of model for explainable recommendation (e.g., matrix factorization, topic modeling, deep learning, etc.) and the information/style of the generated explanation (e.g., textual sentence explanation, etc.). Note that due to the table space this is an incomplete enumeration of the existing explainable recommendation methods, and more methods are introduced in detail in the following parts of the survey. Besides, some of the table cells are empty because to the best of our knowledge there has not been a work falling into the corresponding combination

Information/ style of the explanations	Methods for Explainable Recommendation							
	Neighbor-based	Matrix factorization	Topic modeling	Graph-based	Deep learning	Knowledge-based	Rule mining	Post-hoc
Relevant user or item	Herlocker *et al.*, 2000	Abdollahi and Nasraoui, 2017		Heckel *et al.*, 2017	Chen *et al.*, 2018c	Catherine *et al.*, 2017	Peake and Wang 2018	Cheng *et al.*, 2019a
User or item features	Vig *et al.*, 2009	Zhang *et al.*, 2014a	McAuley and Leskovec, 2013	He *et al.*, 2015	Seo *et al.*, 2017	Huang *et al.*, 2018	Davidson *et al.*, 2010	McInerney *et al.*, 2018
Textual sentence explanation		Zhang *et al.*, 2014a			Li *et al.*, 2017	Ai *et al.*, 2018	Balog *et al.*, 2019	Wang *et al.*, 2018d
Visual explanation					Chen *et al.*, 2019b			
Social explanation	Sharma and Cosley, 2013		Ren *et al.*, 2017	Park *et al.*, 2018				
Word cluster		Zhang, 2015	Wu and Ester 2015					

research according to this taxonomy, so that readers can understand the relationship between existing explainable recommendation methods.

Due to the large body of related work, Table 1.1 is only an incomplete enumeration of explainable recommendation methods. For each "model – information" combination, we present one representative work in the corresponding table cell. However, in Sections 2 and 3 of the survey, we will introduce the details of many explainable recommendation methods.

1.4 Explainability and Effectiveness

Explainability and effectiveness could sometimes be conflicting goals in model design that we have to trade-off (Ricci *et al.*, 2011), i.e., we can either choose a simple model for better explainability, or choose a complex model for better accuracy while sacrificing the explainability. While recent evidence also suggests that these two goals may not necessarily conflict with each other when designing recommendation models (Bilgic *et al.*, 2004; Zhang *et al.*, 2014a). For example, state-of-the-art techniques – such as the deep representation learning approaches – can help us to design recommendation models that are both effective and explainable. Developing explainable deep models is also an attractive direction in the broader AI community, leading to progress not only in explainable recommendation research, but also in fundamental explainable machine learning problems.

When introducing each explainable recommendation model in the following sections, we will also discuss the relationship between explainability and effectiveness in personalized recommendations.

1.5 Explainability and Interpretability

Explainability and interpretability are closely related concepts in the literature. In general, interpretability is one of the approaches to achieve explainability. More specifically, Explainable AI (XAI) aims to develop models that can explain their (or other model's) decisions for system designers or normal users. To achieve the goal, the model can be either interpretable or non-interpretable. For example, interpretable models (such as interpretable machine learning) try to develop models whose decision mechanism is locally or globally transparent, and in this way, the model outputs are usually naturally explainable. Prominent examples of interpretable models include many linear models such as linear regression and tree-based models such as decision trees. Meanwhile, interpretability is not the only way to achieve explainability, e.g., some models can reveal their internal decision mechanism for explanation purpose with complex explanation techniques, such as neural attention mechanisms, natural language explanations, and many post-hoc explanation models, which are widely used in information retrieval, natural language processing, computer vision, graph analysis, and many other tasks. Researchers and practitioners may design and select appropriate explanation methods to achieve explainable AI for different tasks.

1.6 How to Read the Survey

Potential readers of the survey include both researchers and practitioners interested in explainable recommendation systems. Readers are encouraged to prepare with basic understandings of recommender systems, such as content-based recommendation (Pazzani and Billsus, 2007), collaborative filtering (Ekstrand *et al.*, 2011), and evaluation of recommender systems (Shani and Gunawardana, 2011). It is also beneficial to read other related surveys such as explanations in recommender systems from a user study perspective (Tintarev and Masthoff, 2007a), interpretable machine learning (Lipton, 2018; Molnar, 2019), as well as explainable AI in general (Gunning, 2017; Samek *et al.*, 2017).

The following part of the survey will be organized as follows. In Section 2 we will review explainable recommendation from a user-interaction perspective. Specifically, we will discuss different information sources that can facilitate explainable recommendation, and different display styles of recommendation explanation, which are closely related with the corresponding information source. Section 3 will focus on a machine learning perspective of explainable recommendation, which will introduce different types of models for explainable recommendation. Section 4 will introduce evaluation protocols for explainable recommendation, while Section 5 introduces how explainable recommendation methods are used in different real-world recommender system applications. In Section 6 we will summarize the survey with several important open problems and future directions of explainable recommendation research.

References

Abdollahi, B. and O. Nasraoui (2016). “Explainable matrix factorization for collaborative filtering”. In: *Proceedings of the 25th International Conference Companion on World Wide Web*. International World Wide Web Conferences Steering Committee. 5–6.

Abdollahi, B. and O. Nasraoui (2017). “Using explainability for constrained matrix factorization”. In: *Proceedings of the 11th ACM Conference on Recommender Systems*. ACM. 79–83.

Adomavicius, G. and A. Tuzhilin (2011). “Context-aware recommender systems”. In: *Recommender Systems Handbook*. Springer. 217–253.

Agarwal, R., R. Srikant, *et al.* (1994). “Fast algorithms for mining association rules”. In: *Proceedings of the 20th VLDB Conference*. 487–499.

Agrawal, R., T. Imieliński, and A. Swami (1993). “Mining association rules between sets of items in large databases”. *ACM SIGMOD Record*. 22(2): 207–216.

Ai, Q., V. Azizi, X. Chen, and Y. Zhang (2018). “Learning heterogeneous knowledge base embeddings for explainable recommendation”. *Algorithms*. 11(9): 137.

Ai, Q., Y. Zhang, K. Bi, and W. B. Croft (2019). “Explainable product search with a dynamic relation embedding model”. *ACM Transactions on Information Systems (TOIS)*. 38(1): 4.

Amatriain, X. and J. M. Pujol (2015). "Data mining methods for recommender systems". In: *Recommender Systems Handbook*. Springer. 227–262.

Arnott, D. (2006). "Cognitive biases and decision support systems development: A design science approach". *Information Systems Journal*. 16(1): 55–78.

Balabanović, M. and Y. Shoham (1997). "Fab: Content-based, collaborative recommendation". *Communications of the ACM*. 40(3): 66–72.

Balog, K., F. Radlinski, and S. Arakelyan (2019). "Transparent, scrutable and explainable user models for personalized recommendation". In: *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM.

Baral, R., X. Zhu, S. Iyengar, and T. Li (2018). "ReEL: Review aware explanation of location recommendation". In: *Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization*. ACM. 23–32.

Bauman, K., B. Liu, and A. Tuzhilin (2017). "Aspect based recommendations: Recommending items with the most valuable aspects based on user reviews". In: *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM. 717–725.

Bilgic, M., R. Mooney, and E. Rich (2004). "Explanation for recommender systems: Satisfaction vs. promotion". Computer Sciences Austin, University of Texas. Undergraduate Honors. 27.

Bilgic, M. and R. J. Mooney (2005). "Explaining recommendations: Satisfaction vs. promotion". In: *Proceedings of Beyond Personalization 2005: A Workshop on the Next Stage of Recommender Systems Research at the 2005 International Conference on Intelligent User Interfaces, San Diego, CA, USA*.

Bountouridis, D., M. Marrero, N. Tintarev, and C. Hauff (2018). "Explaining credibility in news articles using cross-referencing". In: *Proceedings of the 1st International Workshop on ExplainAble Recommendation and Search (EARS 2018). Ann Arbor, MI, USA*.

Burke, K. and S. Leben (2007). "Procedural fairness: A key ingredient in public satisfaction". *Court Review*. 44(1-2): 4–25.

Catherine, R., K. Mazaitis, M. Eskenazi, and W. Cohen (2017). “Explainable entity-based recommendations with knowledge graphs”. In: *Proceedings of the Poster Track of the 11th ACM Conference on Recommender Systems.* ACM.

Celma, O. (2010). “Music recommendation”. In: *Music Recommendation and Discovery*. Springer. 43–85.

Chaney, A. J., D. M. Blei, and T. Eliassi-Rad (2015). “A probabilistic model for using social networks in personalized item recommendation”. In: *Proceedings of the 9th ACM Conference on Recommender Systems*. ACM. 43–50.

Chang, S., F. M. Harper, and L. G. Terveen (2016). “Crowd-based personalized natural language explanations for recommendations”. In: *Proceedings of the 10th ACM Conference on Recommender Systems*. ACM. 175–182.

Chen, C., M. Zhang, Y. Liu, and S. Ma (2018a). “Neural attentional rating regression with review-level explanations”. In: *Proceedings of the 2018 World Wide Web Conference*. 1583–1592.

Chen, H., X. Chen, S. Shi, and Y. Zhang (2019a). “Generate natural language explanations for recommendation”. In: *Proceedings of the SIGIR 2019 Workshop on ExplainAble Recommendation and Search (EARS 2019)*.

Chen, H., S. Shi, Y. Li, and Y. Zhang (2020). “From collaborative filtering to collaborative reasoning”. *arXiv preprint arXiv:1910.08629*.

Chen, J., F. Zhuang, X. Hong, X. Ao, X. Xie, and Q. He (2018b). “Attention-driven factor model for explainable personalized recommendation”. In: *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. ACM. 909–912.

Chen, X., H. Chen, H. Xu, Y. Zhang, Y. Cao, Z. Qin, and H. Zha (2019b). “Personalized fashion recommendation with visual explanations based on multimodal attention network: Towards visually explainable recommendation”. In: *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. 765–774.

Chen, X., Z. Qin, Y. Zhang, and T. Xu (2016). "Learning to rank features for recommendation over multiple categories". In: *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*. ACM. 305–314.

Chen, X., H. Xu, Y. Zhang, Y. Cao, H. Zha, Z. Qin, and J. Tang (2018c). "Sequential recommendation with user memory networks". In: *Proceedings of the 11th ACM International Conference on Web Search and Data Mining*. ACM.

Chen, X., Y. Zhang, and Z. Qin (2019c). "Dynamic explainable recommendation based on neural attentive models". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. AAAI. 53–60.

Chen, Z., X. Wang, X. Xie, T. Wu, G. Bu, Y. Wang, and E. Chen (2019d). "Co-attentive multi-task learning for explainable recommendation". In: *Proceedings of the 28th International Joint Conference on Artificial Intelligence*. IJCAI. 2137–2143.

Cheng, H.-T., L. Koc, J. Harmsen, T. Shaked, T. Chandra, H. Aradhye, G. Anderson, G. Corrado, W. Chai, M. Ispir, *et al.* (2016). "Wide & deep learning for recommender systems". In: *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems*. ACM. 7–10.

Cheng, W., Y. Shen, L. Huang, and Y. Zhu (2019a). "Incorporating interpretability into latent factor models via fast influence analysis". In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM. 885–893.

Cheng, Z., X. Chang, L. Zhu, R. C. Kanjirathinkal, and M. Kankanhalli (2019b). "MMALFM: Explainable recommendation by leveraging reviews and images". *ACM Transactions on Information Systems (TOIS)*. 37(2): 16.

Cho, Y. H., J. K. Kim, and S. H. Kim (2002). "A personalized recommender system based on web usage mining and decision tree induction". *Expert systems with Applications*. 23(3): 329–342.

Clancey, W. J. (1982). "The epistemology of a rule-based expert system: A framework for explanation." *Tech. Rep.* Department of Computer Science, Stanford University, CA.

Cleger-Tamayo, S., J. M. Fernandez-Luna, and J. F. Huete (2012). “Explaining neighborhood-based recommendations”. In: *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. 1063–1064.

Costa, F., S. Ouyang, P. Dolog, and A. Lawlor (2018). “Automatic generation of natural language explanations”. In: *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion*. ACM. 57.

Covington, P., J. Adams, and E. Sargin (2016). “Deep neural networks for YouTube recommendations”. In: *Proceedings of the 10th ACM Conference on Recommender Systems*. ACM. 191–198.

Cowgill, B. and C. E. Tucker (2019). “Economics, fairness and algorithmic bias”. Columbia Business School Research Paper. Forthcoming.

Cramer, H., V. Evers, S. Ramlal, M. Van Someren, L. Rutledge, N. Stash, L. Aroyo, and B. Wielinga (2008). “The effects of transparency on trust in and acceptance of a content-based art recommender”. *User Modeling and User-Adapted Interaction*. 18(5): 455.

Dacrema, M. F., P. Cremonesi, and D. Jannach (2019). “Are we really making much progress? A worrying analysis of recent neural recommendation approaches”. In: *Proceedings of the 13th ACM Conference on Recommender Systems*. ACM. 101–109.

Davidson, J., B. Liebald, J. Liu, P. Nandy, T. Van Vleet, U. Gargi, S. Gupta, Y. He, M. Lambert, B. Livingston, *et al.* (2010). “The YouTube video recommendation system”. In: *Proceedings of the 4th ACM conference on Recommender systems*. ACM. 293–296.

Devlin, J., M.-W. Chang, K. Lee, and K. Toutanova (2018). “Bert: Pre-training of deep bidirectional transformers for language understanding”. *arXiv preprint arXiv:1810.04805*.

Dietvorst, B. J., J. P. Simmons, and C. Massey (2015). “Algorithm aversion: People erroneously avoid algorithms after seeing them err”. *Journal of Experimental Psychology: General*. 144(1): 114.

Donkers, T., B. Loepp, and J. Ziegler (2017). “Sequential user-based recurrent neural network recommendations”. In: *Proceedings of the 11th ACM Conference on Recommender Systems*. ACM. 152–160.

Du, F., C. Plaisant, N. Spring, K. Crowley, and B. Shneiderman (2019). "EventAction: A visual analytics approach to explainable recommendation for event sequences". *ACM Transactions on Interactive Intelligent Systems (TiiS)*. 9(4): 21.

Ekstrand, M. D. *et al.* (2011). "Collaborative filtering recommender systems". *Foundations and Trends® in Human–Computer Interaction*. 4(2): 81–173.

Ferwerda, B., K. Swelsen, and E. Yang (2012). "Explaining content-based recommendations". New York. 1–24.

Fine Licht, J. de (2014). "Transparency actually: How transparency affects public perceptions of political decision-making". *European Political Science Review*. 6(2): 309–330.

Flesch, R. (1948). "A new readability yardstick". *Journal of Applied Psychology*. 32(3): 221.

Gao, J., X. Wang, Y. Wang, and X. Xie (2019). "Explainable recommendation through attentive multi-view learning". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. AAAI.

Gao, J., N. Liu, M. Lawley, and X. Hu (2017). "An interpretable classification framework for information extraction from online healthcare forums". *Journal of Healthcare Engineering*. 2017.

Gunning, D. (2017). "Explainable artificial intelligence (XAI)". *Defense Advanced Research Projects Agency (DARPA)*.

Gunning, R. (1952). "The technique of clear writing". *Information Transfer and Management*. McGraw-Hill.

He, X., T. Chen, M.-Y. Kan, and X. Chen (2015). "Trirank: Review-aware explainable recommendation by modeling aspects". In: *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*. ACM. 1661–1670.

Heckel, R., M. Vlachos, T. Parnell, and C. Duenner (2017). "Scalable and interpretable product recommendations via overlapping co-clustering". In: *2017 IEEE 33rd International Conference on Data Engineering (ICDE)*. IEEE. 1033–1044.

Herlocker, J. L., J. A. Konstan, and J. Riedl (2000). "Explaining collaborative filtering recommendations". In: *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*. ACM. 241–250.

Herlocker, J. L. and J. A. Konstan (2000). *Understanding and Improving Automated Collaborative Filtering Systems*. University of Minnesota Minnesota.

Hidasi, B., A. Karatzoglou, L. Baltrunas, and D. Tikk (2016). "Session-based recommendations with recurrent neural networks". In: *Proceedings of the 4th International Conference on Learning Representations*.

Hou, Y., N. Yang, Y. Wu, and S. Y. Philip (2018). "Explainable recommendation with fusion of aspect information". *World Wide Web*. 1–20.

Hu, M. and B. Liu (2004). "Mining and summarizing customer reviews". In: *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM. 168–177.

Huang, J., W. X. Zhao, H. Dou, J.-R. Wen, and E. Y. Chang (2018). "Improving sequential recommendation with knowledge-enhanced memory networks". In: *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. ACM. 505–514.

Huang, X., Q. Fang, S. Qian, J. Sang, Y. Li, and C. Xu (2019). "Explainable interaction-driven user modeling over knowledge graph for sequential recommendation". In: *Proceedings of the 27th ACM International Conference on Multimedia*. ACM. 548–556.

Jain, S. and B. C. Wallace (2019). "Attention is not explanation". In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Association for Computational Linguistics. 3543–3556.

Kincaid, J. P., R. P. Fishburne Jr, R. L. Rogers, and B. S. Chissom (1975). "Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel". *Tech. Rep.* Naval Technical Training Command Millington TN Research Branch.

Kittur, A., E. H. Chi, and B. Suh (2008). "Crowdsourcing user studies with mechanical turk". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 453–456.

Knijnenburg, B. P., M. C. Willemsen, Z. Gantner, H. Soncu, and C. Newell (2012). "Explaining the user experience of recommender systems". *User Modeling and User-Adapted Interaction*. 22(4-5): 441–504.

Koh, P. W. and P. Liang (2017). "Understanding black-box predictions via influence functions". In: *Proceedings of the 34th International Conference on Machine Learning–Volume 70*. JMLR. 1885–1894.

Koren, Y. (2008). "Factorization meets the neighborhood: A multi-faceted collaborative filtering model". In: *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM. 426–434.

Koren, Y., R. Bell, and C. Volinsky (2009). "Matrix factorization techniques for recommender systems". *Computer*. 42(8): 42–49.

Kraus, C. L. (2016). "A news recommendation engine for a multi-perspective understanding of political topics". *Master Thesis, Technical University of Berlin*.

Lee, D. D. and H. S. Seung (1999). "Learning the parts of objects by non-negative matrix factorization". *Nature*. 401(6755): 788.

Lee, D. D. and H. S. Seung (2001). "Algorithms for non-negative matrix factorization". *Advances in Neural Information Processing Systems*. 13: 556–562.

Lee, O.-J. and J. J. Jung (2018). "Explainable movie recommendation systems by using story-based similarity". In: *Proceedings of the ACM IUI 2018 Workshops*. ACM IUI.

Li, C., C. Quan, L. Peng, Y. Qi, Y. Deng, and L. Wu (2019). "A capsule network for recommendation and explaining what you like and dislike". In: *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. 275–284.

Li, P., Z. Wang, Z. Ren, L. Bing, and W. Lam (2017). "Neural rating regression with abstractive tips generation for recommendation". In: *Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval*. ACM. 345–354.

Lin, C.-Y. (2004). "Rouge: A package for automatic evaluation of summaries". In: *Text Summarization Branches Out*. ACL.

Lin, W., S. A. Alvarez, and C. Ruiz (2000). "Collaborative recommendation via adaptive association rule mining". In: *Proceedings of the International Workshop on Web Mining for E-Commerce (WEBKDD'2000).* ACM.

Lin, W., S. A. Alvarez, and C. Ruiz (2002). "Efficient adaptive-support association rule mining for recommender systems". *Data Mining and Knowledge Discovery*. 6(1): 83–105.

Lin, Y., P. Ren, Z. Chen, Z. Ren, J. Ma, and M. de Rijke (2019). "Explainable fashion recommendation with joint outfit matching and comment generation". *IEEE Transactions on Knowledge and Data Engineering*. 32(4): 1–16.

Linden, G., B. Smith, and J. York (2003). "Amazon.com recommendations: Item-to-item collaborative filtering". *IEEE Internet Computing*. 7(1): 76–80.

Lipton, Z. C. (2018). "The mythos of model interpretability". *Communications of the ACM*. 61(10): 36–43.

Liu, B. (2012). "Sentiment analysis and opinion mining". *Synthesis Lectures on Human Language Technologies*. 5(1): 1–167.

Liu, H., J. Wen, L. Jing, J. Yu, X. Zhang, and M. Zhang (2019). "In2Rec: Influence-based interpretable recommendation". In: *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. ACM. 1803–1812.

Liu, N., D. Shin, and X. Hu (2018). "Contextual outlier interpretation". In: *Proceedings of the 27th International Joint Conference on Artificial Intelligence*. IJCAI. 2461–2467.

Lu, Y., R. Dong, and B. Smyth (2018a). "Coevolutionary recommendation model: Mutual learning between ratings and reviews". In: *Proceedings of the 2018 World Wide Web Conference on World Wide Web*. International World Wide Web Conferences Steering Committee. 773–782.

Lu, Y., R. Dong, and B. Smyth (2018b). "Why I like it: Multi-task learning for recommendation and explanation". In: *Proceedings of the 12th ACM Conference on Recommender Systems*. ACM. 4–12.

Lu, Y., M. Castellanos, U. Dayal, and C. Zhai (2011). "Automatic construction of a context-aware sentiment lexicon: an optimization approach". In: *Proceedings of the 20th International Conference on World Wide Web*. ACM. 347–356.

Ma, J., C. Zhou, P. Cui, H. Yang, and W. Zhu (2019a). "Learning Disentangled Representations for Recommendation". In: *Advances in Neural Information Processing Systems*. 5712–5723.

Ma, W., M. Zhang, Y. Cao, W. Jin, C. Wang, Y. Liu, S. Ma, and X. Ren (2019b). "Jointly learning explainable rules for recommendation with knowledge graph". In: *The World Wide Web Conference*. ACM. 1210–1221.

Mc Laughlin, G. H. (1969). "SMOG grading-a new readability formula". *Journal of Reading*. 12(8): 639–646.

McAuley, J. and J. Leskovec (2013). "Hidden factors and hidden topics: understanding rating dimensions with review text". In: *Proceedings of the 7th ACM Conference on Recommender Systems*. ACM. 165–172.

McInerney, J., B. Lacker, S. Hansen, K. Higley, H. Bouchard, A. Gruson, and R. Mehrotra (2018). "Explore, exploit, and explain: Personalizing explainable recommendations with bandits". In: *Proceedings of the 12th ACM Conference on Recommender Systems*. ACM. 31–39.

McSherry, D. (2005). "Explanation in recommender systems". *Artificial Intelligence Review*. 24(2): 179–197.

Miller, T. (2019). "Explanation in artificial intelligence: Insights from the social sciences". *Artificial Intelligence*. 267: 1–38.

Mnih, A. and R. R. Salakhutdinov (2008). "Probabilistic matrix factorization". In: *Advances in Neural Information Processing Systems*. 1257–1264.

Mobasher, B., H. Dai, T. Luo, and M. Nakagawa (2001). "Effective personalization based on association rule discovery from web usage data". In: *Proceedings of the 3rd International Workshop on Web Information and Data Management*. ACM. 9–15.

Molnar, C. (2019). *Interpretable Machine Learning*. Leanpub.

Nanou, T., G. Lekakos, and K. Fouskas (2010). "The effects of recommendations? presentation on persuasion and satisfaction in a movie recommender system". *Multimedia Systems*. 16(4-5): 219–230.

Ni, J., J. Li, and J. 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)*. 188–197.

Papadimitriou, A., P. Symeonidis, and Y. Manolopoulos (2012). "A generalized taxonomy of explanations styles for traditional and social recommender systems". *Data Mining and Knowledge Discovery*. 24(3): 555–583.

Papineni, K., S. Roukos, T. Ward, and W.-J. Zhu (2002). "BLEU: a method for automatic evaluation of machine translation". In: *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*. Association for Computational Linguistics. 311–318.

Park, H., H. Jeon, J. Kim, B. Ahn, and U. Kang (2018). "UniWalk: Explainable and accurate recommendation for rating and network data". *arXiv preprint arXiv:1710.07134*.

Pazzani, M. J. (1999). "A framework for collaborative, content-based and demographic filtering". *Artificial Intelligence Review*. 13(5-6): 393–408.

Pazzani, M. J. and D. Billsus (2007). "Content-based recommendation systems". In: *The Adaptive Web*. Springer. 325–341.

Peake, G. and J. Wang (2018). "Explanation mining: Post hoc interpretability of latent factor models for recommendation systems". In: *Proceedings of Beyond Personalization 2005: A Workshop on the Next Stage of Recommender Systems Research at the 2005 International Conference on Intelligent User Interfaces, San Diego, CA, USA*. ACM. 2060–2069.

Pei, K., Y. Cao, J. Yang, and S. Jana (2017). "Deepxplore: Automated whitebox testing of deep learning systems". In: *Proceedings of the 26th Symposium on Operating Systems Principles*. ACM. 1–18.

Qiu, L., S. Gao, W. Cheng, and J. Guo (2016). "Aspect-based latent factor model by integrating ratings and reviews for recommender system". *Knowledge-Based Systems*. 110: 233–243.

Quijano-Sanchez, L., C. Sauer, J. A. Recio-Garcia, and B. Diaz-Agudo (2017). "Make it personal: A social explanation system applied to group recommendations". *Expert Systems with Applications*. 76: 36–48.

Ren, Z., S. Liang, P. Li, S. Wang, and M. de Rijke (2017). "Social collaborative viewpoint regression with explainable recommendations". In: *Proceedings of the 10th ACM International Conference on Web Search and Data Mining*. ACM. 485–494.

Rendle, S., C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme (2009). "BPR: Bayesian personalized ranking from implicit feedback". In: *Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence*. AUAI Press. 452–461.

Rendle, S. and L. Schmidt-Thieme (2010). "Pairwise interaction tensor factorization for personalized tag recommendation". In: *Proceedings of the 3rd ACM International Conference on Web Search and Data Mining*. ACM. 81–90.

Rennie, J. D. and N. Srebro (2005). "Fast maximum margin matrix factorization for collaborative prediction". In: *Proceedings of the 22nd International Conference on Machine Learning*. ACM. 713–719.

Resnick, P., N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl (1994). "GroupLens: An open architecture for collaborative filtering of netnews". In: *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work*. ACM. 175–186.

Ribeiro, M. T., S. Singh, and C. Guestrin (2016). "Why should I trust you?: Explaining the predictions of any classifier". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM. 1135–1144.

Ricci, F., L. Rokach, and B. Shapira (2011). "Introduction to recommender systems handbook". In: *Recommender Systems Handbook*. Springer. 1–35.

Salakhutdinov, R. and A. Mnih (2008). "Bayesian probabilistic matrix factorization using Markov chain Monte Carlo". In: *Proceedings of the 25th International Conference on Machine Learning*. ACM. 880–887.

Samek, W., T. Wiegand, and K.-R. Müller (2017). "Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models". *arXiv preprint arXiv:1708.08296*.

Sandvig, J. J., B. Mobasher, and R. Burke (2007). "Robustness of collaborative recommendation based on association rule mining". In: *Proceedings of the 2007 ACM Conference on Recommender Systems*. ACM. 105–112.

Sarwar, B., G. Karypis, J. Konstan, and J. Riedl (2001). "Item-based collaborative filtering recommendation algorithms". In: *Proceedings of the 10th International Conference on World Wide Web*. ACM. 285–295.

Schafer, J. B., J. A. Konstan, and J. Riedl (2001). "E-commerce recommendation applications". *Data Mining and Knowledge Discovery*. 5(1-2): 115–153.

Schafer, J. B., J. Konstan, and J. Riedl (1999). "Recommender systems in e-commerce". In: *Proceedings of the 1st ACM Conference on Electronic Commerce*. ACM. 158–166.

Senter, R. and E. A. Smith (1967). "Automated readability index". *Tech. Rep.* Cincinnati University, OH.

Seo, S., J. Huang, H. Yang, and Y. Liu (2017). "Interpretable convolutional neural networks with dual local and global attention for review rating prediction". In: *Proceedings of the 11th ACM Conference on Recommender Systems*. ACM. 297–305.

Shani, G. and A. Gunawardana (2011). "Evaluating recommendation systems". In: *Recommender Systems Handbook*. Springer. 257–297.

Sharma, A. and D. Cosley (2013). "Do social explanations work?: Studying and modeling the effects of social explanations in recommender systems". In: *Proceedings of the 22nd International Conference on World Wide Web*. ACM. 1133–1144.

Sherchan, W., S. Nepal, and C. Paris (2013). "A survey of trust in social networks". *ACM Computing Surveys (CSUR)*. 45(4): 47.

Shi, S., H. Chen, M. Zhang, and Y. Zhang (2019). "Neural logic networks". *arXiv preprint arXiv:1910.08629*.

Shi, Y., M. Larson, and A. Hanjalic (2010). "List-wise learning to rank with matrix factorization for collaborative filtering". In: *Proceedings of the 4th ACM Conference on Recommender Systems*. ACM. 269–272.

Singh, J. and A. Anand (2018). "Posthoc interpretability of learning to rank models using secondary training data". *Proceedings of the SIGIR 2018 International Workshop on ExplainAble Recommendation and Search (EARS)*.

Singh, J. and A. Anand (2019). "EXS: Explainable search using local model agnostic interpretability". In: *Proceedings of the 12th ACM International Conference on Web Search and Data Mining*. ACM. 770–773.

Sinha, R. and K. Swearingen (2002). "The role of transparency in recommender systems". In: *CHI'02 Extended Abstracts on Human Factors in Computing Systems*. ACM. 830–831.

Smyth, B., K. McCarthy, J. Reilly, D. O'Sullivan, L. McGinty, and D. C. Wilson (2005). "Case studies in association rule mining for recommender systems". In: *Proceedings of the 2005 International Conference on Artificial Intelligence*. ICAI. 809–815.

Srebro, N. and T. Jaakkola (2003). "Weighted low-rank approximations". In: *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*. 720–727.

Srebro, N., J. Rennie, and T. S. Jaakkola (2005). "Maximum-margin matrix factorization". In: *Advances in Neural Information Processing Systems*. 1329–1336.

Al-Taie, M. Z. and S. Kadry (2014). "Visualization of explanations in recommender systems". *The Journal of Advanced Management Science*. 2(2): 140–144.

Tan, Y., M. Zhang, Y. Liu, and S. Ma (2016). "Rating-boosted latent topics: Understanding users and items with ratings and reviews". In: *Proceedings of the 25th International Joint Conference on Artificial Intelligence*. IJCAI. 2640–2646.

Tang, J. and K. Wang (2018). "Personalized top-N sequential recommendation via convolutional sequence embedding". In: *Proceedings of the 11th ACM International Conference on Web Search and Data Mining*. ACM.

Tao, Y., Y. Jia, N. Wang, and H. Wang (2019a). “The FacT: Taming latent factor models for explainability with factorization trees”. In: *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM.

Tao, Z., S. Li, Z. Wang, C. Fang, L. Yang, H. Zhao, and Y. Fu (2019b). “Log2Intent: Towards interpretable user modeling via recurrent semantics memory unit”. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM. 1055–1063.

Tintarev, N. (2007). “Explanations of recommendations”. In: *Proceedings of the 2007 ACM Conference on Recommender Systems*. ACM. 203–206.

Tintarev, N. and J. Masthoff (2007a). “A survey of explanations in recommender systems”. In: *Data Engineering Workshop, 2007 IEEE 23rd International Conference*. IEEE. 801–810.

Tintarev, N. and J. Masthoff (2007b). “Effective explanations of recommendations: User-centered design”. In: *Proceedings of the 2007 ACM Conference on Recommender Systems*. ACM. 153–156.

Tintarev, N. and J. Masthoff (2008). “The effectiveness of personalized movie explanations: An experiment using commercial meta-data”. In: *Adaptive Hypermedia and Adaptive Web-Based Systems*. Springer. 204–213.

Tintarev, N. and J. Masthoff (2011). “Designing and evaluating explanations for recommender systems”. In: *Recommender Systems Handbook*. Springer. 479–510.

Tintarev, N. and J. Masthoff (2015). “Explaining recommendations: Design and evaluation”. In: *Recommender Systems Handbook*. Springer. 353–382.

Toderici, G., H. Aradhye, M. Pasca, L. Sbaiz, and J. Yagnik (2010). “Finding meaning on YouTube: Tag recommendation and category discovery”. In: *The 23rd IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2010*. IEEE. 3447–3454.

Tsai, C.-H. and P. Brusilovsky (2018). “Explaining social recommendations to casual users: Design principles and opportunities”. In: *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion*. ACM. 59.

Tsukuda, K. and M. Goto (2019). “DualDiv: Diversifying items and explanation styles in explainable hybrid recommendation”. In: *Proceedings of the 13th ACM Conference on Recommender Systems*. ACM. 398–402.

Vig, J., S. Sen, and J. Riedl (2009). “Tagsplanations: Explaining recommendations using tags”. In: *Proceedings of the 14th International Conference on Intelligent User Interfaces*. ACM. 47–56.

Wang, B., M. Ester, J. Bu, and D. Cai (2014). “Who Also Likes It? Generating the Most Persuasive Social Explanations in Recommender Systems.” In: *Proceedings of the AAAI Conference on Artificial Intelligence*. AAAI. 173–179.

Wang, H., F. Zhang, J. Wang, M. Zhao, W. Li, X. Xie, and M. Guo (2018a). “RippleNet: Propagating user preferences on the knowledge graph for recommender systems”. In: *Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM 2018)*. ACM. 417–426.

Wang, N., H. Wang, Y. Jia, and Y. Yin (2018b). “Explainable recommendation via multi-task learning in opinionated text data”. In: *Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. ACM.

Wang, W. and I. Benbasat (2007). “Recommendation agents for electronic commerce: Effects of explanation facilities on trusting beliefs”. *Journal of Management Information Systems*. 23(4): 217–246.

Wang, X., X. He, F. Feng, L. Nie, and T.-S. Chua (2018c). “TEM: Tree-enhanced embedding model for explainable recommendation”. In: *Proceedings of the 27th International Conference on World Wide Web. International World Wide Web Conferences Steering Committee*.

Wang, X., Y. Chen, J. Yang, L. Wu, Z. Wu, and X. Xie (2018d). “A reinforcement learning framework for explainable recommendation”. In: *2018 IEEE International Conference on Data Mining (ICDM)*. IEEE. 587–596.

Wiegreffe, S. and Y. Pinter (2019). “Attention is not not explanation”. 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. 11–20.

Wu, L., C. Quan, C. Li, Q. Wang, B. Zheng, and X. Luo (2019). “A context-aware user-item representation learning for item recommendation”. *ACM Transactions on Information Systems (TOIS)*. 37(2): 22.

Wu, Y., C. DuBois, A. X. Zheng, and M. Ester (2016). “Collaborative denoising auto-encoders for top-n recommender systems”. In: *Proceedings of the 9th ACM International Conference on Web Search and Data Mining*. ACM. 153–162.

Wu, Y. and M. Ester (2015). “Flame: A probabilistic model combining aspect based opinion mining and collaborative filtering”. In: *Proceedings of the 8th ACM International Conference on Web Search and Data Mining*. ACM. 199–208.

Xian, Y., Z. Fu, S. Muthukrishnan, G. de Melo, and Y. Zhang (2019). “Reinforcement knowledge graph reasoning for explainable recommendation”. In: *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM.

Yu, C., L. V. Lakshmanan, and S. Amer-Yahia (2009). “Recommendation diversification using explanations”. In: *Data Engineering, 2009. ICDE’09. IEEE 25th International Conference*. IEEE. 1299–1302.

Zanker, M. and D. Ninaus (2010). “Knowledgeable explanations for recommender systems”. In: *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference*. IEEE. 657–660.

Zhang, S., L. Yao, A. Sun, and Y. Tay (2019). “Deep learning based recommender system: A survey and new perspectives”. *ACM Computing Surveys (CSUR)*. 52(1): 5.

Zhang, Y. (2015). “Incorporating phrase-level sentiment analysis on textual reviews for personalized recommendation”. In: *Proceedings of the 8th ACM International Conference on Web Search and Data Mining*. ACM. 435–440.

Zhang, Y., G. Lai, M. Zhang, Y. Zhang, Y. Liu, and S. Ma (2014a). "Explicit factor models for explainable recommendation based on phrase-level sentiment analysis". In: *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*. ACM. 83–92.

Zhang, Y., H. Zhang, M. Zhang, Y. Liu, and S. Ma (2014b). "Do users rate or review?: Boost phrase-level sentiment labeling with review-level sentiment classification". In: *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*. ACM. 1027–1030.

Zhang, Y., M. Zhang, Y. Liu, and S. Ma (2013a). "Improve collaborative filtering through bordered block diagonal form matrices". In: *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. 313–322.

Zhang, Y., M. Zhang, Y. Liu, S. Ma, and S. Feng (2013b). "Localized matrix factorization for recommendation based on matrix block diagonal forms". In: *Proceedings of the 22nd International Conference on World Wide Web*. ACM. 1511–1520.

Zhang, Y., M. Zhang, Y. Liu, C. Tat-Seng, Y. Zhang, and S. Ma (2015a). "Task-based recommendation on a web-scale". In: *Big Data (Big Data), 2015 IEEE International Conference*. IEEE. 827–836.

Zhang, Y., M. Zhang, Y. Zhang, G. Lai, Y. Liu, H. Zhang, and S. Ma (2015b). "Daily-aware personalized recommendation based on feature-level time series analysis". In: *Proceedings of the 24th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee. 1373–1383.

Zhang, Y., X. Xu, H. Zhou, and Y. Zhang (2020). "Distilling structured knowledge into embeddings for explainable and accurate recommendation". In: *Proceedings of the 13th ACM International Conference on Web Search and Data Mining (WSDM)*. ACM.

Zhao, G., H. Fu, R. Song, T. Sakai, Z. Chen, X. Xie, and X. Qian (2019a). "Personalized reason generation for explainable song recommendation". *ACM Transactions on Intelligent Systems and Technology (TIST)*. 10(4): 41.

Zhao, J., Z. Guan, and H. Sun (2019b). "Riker: Mining rich keyword representations for interpretable product question answering". In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM. 1389–1398.

Zhao, K., G. Cong, Q. Yuan, and K. Q. Zhu (2015). "SAR: A sentiment-aspect-region model for user preference analysis in geo-tagged reviews". In: *Data Engineering (ICDE), 2015 IEEE 31st International Conference*. IEEE. 675–686.

Zhao, W. X., S. Li, Y. He, L. Wang, J.-R. Wen, and X. Li (2016). "Exploring demographic information in social media for product recommendation". *Knowledge and Information Systems*. 49(1): 61–89.

Zhao, X. W., Y. Guo, Y. He, H. Jiang, Y. Wu, and X. Li (2014). "We know what you want to buy: a demographic-based system for product recommendation on microblogs". In: *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM. 1935–1944.

Zheng, L., V. Noroozi, and P. S. Yu (2017). "Joint deep modeling of users and items using reviews for recommendation". In: *Proceedings of the 10th ACM International Conference on Web Search and Data Mining*. ACM. 425–434.