Transparency, Privacy, and Fairness
in
Recommender Systems

SUBMITTED BY

DOMINIK KOWALD

Institute of Interactive Systems and Data Science
Graz University of Technology

Graz University of Technology

I've got to keep going, be strong. Must be so determined and push myself on.

— Iron Maiden, The Loneliness of the Long Distance Runner —

Dedicated to my wife, Tea.

Overview of Chapters

Chapter 1 - Introduction

This chapter starts with the motivation and the scientific positioning of this habilitation within the broad research field of recommender systems. It also lists and briefly outlines the 17 main publications of this habilitation.

Chapter 2 - Related Work and Background

Chapter 2 briefly discusses the related work and background relevant to this habilitation, namely (i) main concepts of recommender systems, (ii) transparency and cognitive models in recommender systems, (iii) privacy and limited preference information in recommender systems, and (iv) fairness and popularity bias in recommender systems. Additionally, this chapter briefly summarizes the author's own research efforts in relation to the related work.

Chapter 3 - Scientific Contributions

This chapter describes the 7 scientific contributions of this habilitation, which are (i) using cognitive models for a transparent design and implementation process of recommender systems, (ii) illustrating to what extent components of the cognitive model ACT-R contribute to recommendations, (iii) addressing limited user preference information in cold-start and session-based recommendation settings, (iv) addressing users' privacy constraints and the trade-off between accuracy and privacy in recommendations, (v) measuring popularity bias for user groups differing in mainstreaminess and gender, (vi) understanding popularity bias mitigation and amplification, and (vii) studying long-term dynamics of fairness in algorithmic decision support. Additionally, this chapter discusses reproducibility aspects of the presented research results and findings.

Chapter 4 - Outlook and Future Research

Chapter 4 gives an outlook into future research directions based on the results, scientific contributions, and findings of this habilitation.

Please note that this is a slightly adapted and updated version of this habilitation reflecting the state of the research conducted until mid 2024. Furthermore, this version of the habilitation does not contain the full texts of the publications, but the DOI links to the online versions are provided in Section 1.2.

Contents

Acknowledgements

Abstract

Recommender systems have become a pervasive part of our daily online experience by analyzing past usage behavior to suggest potential relevant content, e.g., music, movies, or books. Today, recommender systems are one of the most widely used applications of artificial intelligence and machine learning. Therefore, regulations and requirements for trustworthy artificial intelligence, for example, the European AI Act, which includes notions such as transparency, privacy, and fairness are also highly relevant for the design, development, evaluation, and deployment of recommender systems in practice. This habilitation elaborates on aspects related to these three notions in the light of recommender systems, namely: (i) transparency and cognitive models, (ii) privacy and limited preference information, and (iii) fairness and popularity bias in recommender systems.

Specifically, with respect to aspect (i), we highlight the usefulness of incorporating psychological theories for a transparent design process of recommender systems. We term this type of systems psychology-informed recommender systems. We also use models of human memory theory to develop cognitive-inspired algorithms for tag and music recommendations, and find that these algorithms are capable of outperforming related methods in terms of recommendation accuracy. Additionally, we show that cognitive models can further contribute to transparency aspects of recommender systems by illustrating how the models' components have contributed to generate the recommendation lists.

In aspect (ii), we study and address the trade-off between accuracy and privacy in differentially-private recommendations. We design a novel recommendation approach for collaborative filtering based on an efficient neighborhood reuse concept, which reduces the number of users that need to be protected with differential privacy. Furthermore, we address the related issue of limited availability of user preference information, e.g., click data, in the settings of session-based and cold-start recommendations, by using, e.g., variational autoencoders.

With respect to aspect (iii), we analyze popularity bias in collaborative filtering-based recommender systems. We find that the recommendation frequency of an item is positively correlated with this item's popularity. This also leads to the unfair treatment of users with little interest in popular content, since these users receive worse recommendation accuracy results than users with high interest in popular content. We also find that female users are more strongly affected by the algorithms' amplification of popularity bias. Besides, we present results of an online study on popularity bias mitigation in the field of news article recommendations. Finally, we study long-term fairness dynamics in algorithmic decision support in the labor market using agent-based modeling techniques.

Kurzfassung

Empfehlungssysteme sind zu einem allgegenwärtigen Teil unserer täglichen Online-Erfahrung geworden, indem sie das vergangene Nutzerverhalten analysieren, um relevante Inhalte vorzuschlagen, beispielsweise Musik, Filme oder Bücher. Mittlerweile gehören Empfehlungssysteme zu den am weitesten verbreiteten Anwendungen der künstlichen Intelligenz und des maschinellen Lernens. Daher sind Vorschriften für vertrauenswürdige künstliche Intelligenz, welche Anforderungen wie Transparenz, Datenschutz und Fairness umfassen, für die Entwicklung von Empfehlungssystemen relevant. Diese Habilitation untersucht Empfehlungssysteme in Hinblick auf Aspekte, die mit diesen Anforderungen verknüpft sind, nämlich: (i) Transparenz und kognitive Modelle, (ii) Datenschutz und limitierte Präferenz-Informationen, sowie (iii) Fairness und Popularitätsverzerrungen.

Bezüglich Aspekt (i) zeigen wir den Nutzen von psychologischen Theorien für einen transparenten Designprozess von Empfehlungssystemen. Wir bezeichnen diese als Psychologie-inspirierte Empfehlungssysteme. Zusätzlich verwenden wir Modelle der menschlichen Gedächtnistheorie für die Entwicklung von Empfehlungssystemen und zeigen, dass diese Algorithmen verwandte Methoden, in Bezug auf die Vorhersagegenauigkeit, übertreffen. Darüber hinaus zeigen wir, dass die kognitiven Modelle dazu verwendet werden können, um zu illustrieren, welche Komponenten für die Empfehlungsgenerierung wichtig gewesen sind.

In Hinblick auf Aspekt (ii) untersuchen wir die Beziehung zwischen Genauigkeit und Datenschutz in Empfehlungssystemen, die Differential Privacy verwenden. Wir entwickeln einen neuartigen Empfehlungsalgorithmus, der auf einem effizienten Konzept zur Wiederverwendung von Nachbarschaften im kollaborativen Filtern basiert. Dadurch kann der notwendige Einsatz von Differential Privacy minimiert werden. Darüber hinaus adressieren wir ein damit verwandtes Problem, nämlich das der limitierten Nutzerpräferenz-Informationen, z.B., Klick-Daten, durch die Verwendung von z.B., Variational Autoencodern.

Bezüglich Aspekt (iii) analysieren wir den Einfluss der Popularitätsverzerrung auf die Genauigkeit von Empfehlungssystemen. Wir zeigen, dass Popularität und Empfehlungshäufigkeit positiv korreliert sind, welches auch zur unfairen Behandlung von Nutzern führt, die wenig Interesse an populären Inhalten haben. Diese Nutzer erhalten eine geringere Empfehlungsgenauigkeit als Nutzer, die an populären Inhalten interessiert sind. Darüber hinaus zeigen wir, dass weibliche Benutzer stärker von Popularitätsverzerrungen betroffen sind. Wir präsentieren außerdem Ergebnisse einer Online-Studie zur Minderung des Einflusses von Popularitätsverzerrungen. Abschließend untersuchen wir Langzeiteffekte von Fairness in algorithmischen Entscheidungen mittels agentenbasierter Modellierung.

Chapter 1

Introduction

The present postdoctoral thesis is a cumulative habilitation submitted to Graz University of Technology for the scientific subject *Applied Computer Science*. This habilitation summarizes and discusses scientific publications that have been published between 2018 and 2023, i.e., during the habilitation's author's postdoctoral research. This chapter describes the scientific positioning of this habilitation (Section 1.1), and introduces the 17 main publications that constitute this work (Section 1.2). All publications are peer-reviewed, are already published, and contain a digital object identifier (*DOI*). The publications consist of 7 journal articles, 7 conference proceedings contributions, two workshop post-proceedings book chapters, and one workshop paper. The latter was published via the academic distribution service *arXiv* in accordance with the publishing guidelines of the *Workshop on Transparency and Explainability in Adaptive Systems through User Modeling Grounded in Psychological Theory* co-located with *ACM IUI 2020*.

1.1 Scientific Positioning of this Habilitation

This habilitation investigates the research field of recommender systems in general, and aspects of transparency and cognitive models, privacy and limited preference information, and fairness and popularity bias in recommender systems in particular. The research field of recommender systems makes use of multiple aspects of *Applied Computer Science*, including (but not limited to) data science, user modeling, personalization, machine learning, information retrieval, human computer interaction, computational social science, and trustworthy artificial intelligence.

More concretely, recommender systems can be seen as one of the most widely used instantiations of machine learning and artificial intelligence, and accompany us in our daily online experience. Thus, recommender systems have become an integral part of our digital life for supporting humans in finding relevant information in information spaces that are too big or complex for manual filtering (e.g., [47,126,159,231]). Since the early implementations of recommendation algorithms (e.g., [229,230]), these systems analyze past usage behavior in order to build user models, and to suggest items, or even people in social networks [77,145,164], to individual users or to groups of users (e.g., [189, 190]). To build these user models, different techniques have been employed, including traditional approaches

such as collaborative filtering [81], content-based filtering [181], and hybrid recommendations [45], and more recent approaches based on latent representations (or embeddings) and deep learning [57, 275, 288]. Thus, also different types of data sources are utilized for generating recommendations, e.g., preference information such as ratings, and content features of items (see Section 2.1 for more details on recommender systems in general). Apart from that, recent research has illustrated the multi-stakeholder nature of recommender systems [1, 3]. Thus, not only users are affected by recommendations, but also other stakeholders [121], such as platform operators or item providers (e.g., music artists). Balancing the goals of multiple stakeholders is an active research topic, and further illustrates the far-reaching impact of recommender systems on society [48].

The uptake of recommender systems both in academia and industry [52, 120, 127], as well as their human-centric nature, emphasizes that current regulations and requirements for trustworthy artificial intelligence (AI) are also of high importance for the deployment of recommender systems [69]. Trustworthiness entails multiple notions that have been defined and categorized by the European Commission and institutions in other countries. This has led to different regulations and requirements, for example, the *EU Artificial Intelligence Act* [59], or the *United States Regulatory Development Relating to AI* [269]. Although there are differences between these regulations and requirements, all of them include notions related to transparency, privacy, and fairness in AI. These aspects are also highly relevant under the lens of recommender systems, as indicated by recent related research investigating trustworthy recommender systems [84, 85, 97]. This habilitation contributes to this line of research in the following fields:

Transparency and Cognitive Models in Recommender Systems

One issue of modern recommender systems algorithms based on deep learning techniques (e.g., [275, 288]) is that these approaches are mostly based on principles of artificial intelligence rather than human intelligence. This could lead to non-transparent algorithmic decisions that are hard to understand by the system's users [250]. Apart from methods coming from the fields of explainable AI [198] and explainable recommender systems [263], one way to address this issue is to use theories from psychology to enhance the transparency of recommendation models.

This habilitation uses cognitive models of human memory for a transparent design of recommendation approaches [152, 156, 158, 173, 246]. Specifically, we show that models of human episodic memory and activation processes in human memory can help to create transparent and accurate recommendation models. In this respect, we also illustrate to what extent the components of these models contribute to the generation of the recommendation lists [201]. Finally, we survey and categorize research at the intersection of recommender systems and psychology, which we term *psychology-informed recommender systems* [174].

Privacy and Limited Preference Information in Recommender Systems

Recommender systems need to analyze user preference information to calculate personalized recommendations, which could lead to multiple privacy threats to

users [91]. This includes the inference of users' sensitive information (e.g., gender), or the disclosure of users' preference information (e.g., who bought what) via the analysis of generated recommendation lists by untrusted third parties (e.g., [33, 49, 287]). Thus, privacy has become a key requirement for personalized recommender systems, especially in the light of current data protection initiatives such as the *European General Data Protection Regulation (GDPR)*. Therefore, privacy is related to the issue of limited availability of user preference information (e.g., clicks or ratings) due to the restricted utilization of users' preference information as a result of data protection initiatives [40, 62], and due to the increased privacy concerns of users (e.g., users are not willing to share preference information or to sign in to the system) [138, 169, 194]. This could lead to the user cold-start problem [241] and session-based recommendation settings, since long-term user preferences (including past preferences of the target user) are unavailable [125].

This habilitation investigates issues of limited preference information by addressing the user cold-start problem using recommendations based on users' trust connections [73], and by studying the usefulness of variational autoencoders for session-based job recommendations [167]. Additionally, we study varying privacy constraints of users in a matrix factorization-based recommender system using *meta learning* [202]. We also address the privacy-accuracy trade-off in differentially private recommender systems by utilizing an efficient neighborhood reuse concept [206]. Finally, we survey and categorize the literature on employing *differential privacy* for collaborative filtering recommender systems [205].

Fairness and Popularity Bias in Recommender Systems

Although bias and fairness in algorithmic decision support and machine learning is a research topic that has gained a lot of attraction in recent years [37, 161, 195], the reflection and replication of biases is still an open research problem in the field of interactive systems in general [92, 168], and recommender systems in particular [55, 184, 197]. Here, especially popularity bias is a common issue in recommender systems based on collaborative filtering, and leads to the underrepresentation of unpopular content in personalized recommendation lists [9, 20, 83].

The research presented in this habilitation shows that this popularity bias unfairly affects users with little interest in popular content, since this user group receives lower recommendation accuracy than users interested in popular content [147, 153, 154, 157]. We also find that recommendation algorithms could amplify popularity bias for female users [171], and that content-based recommendations can help to mitigate popularity bias [163]. Additionally, we study long-term fairness in algorithmic decision support in the labor market, and find that there is a trade-off between *individual* and *group fairness* in this setting [243].

Reproducibility Aspects of this Habilitation

The reproducibility of recommender systems research results is of utmost importance to be able to track the scientific progress in the field (e.g., [32, 87]). This habilitation provides code and data resources that should foster the reproducibility of the presented research contributions (see Section 3.4 for a full list).

1.2 Main Publications

Table 1.1 lists the 17 main publications of this habilitation. I have selected 5 publications for each of the first two research topics described beforehand. For the third research topic, fairness and popularity bias in recommender systems, I have selected 7 publications, since this is the research topic I have investigated most recently (here, my first paper was published in 2020). Within these three research fields, the publications are sorted by publication year in ascending order. Overall, each publication is assigned a unique ID, i.e., $\boxed{\text{Pi}}$, where $i = 1 \ldots 17$.

In the first field, transparency and cognitive models in recommender systems, the list of publications contains three studies, in which cognitive models are employed for a transparent design process of recommender systems, i.e., one recommendation approach based on a model of human episodic memory $\boxed{P1}$, and two approaches based on models formalizing activation processes in human memory $\boxed{P2}$ $\boxed{P3}$. Furthermore, it lists a survey on psychology-informed recommender systems $\boxed{P4}$. Another publication illustrates to what extent components of cognitive models contribute to the generation of the recommendation lists $\boxed{P5}$.

The second research field contains two studies on addressing the issue of limited availability of user preference information: one addresses the user cold-start problem using trust-based collaborative filtering $\boxed{P6}$, and one employs variational autoencoders for session-based job recommendations $\boxed{P7}$. Table 1.1 also contains three publications on privacy-aware recommender systems, one addressing varying privacy constraints of users $\boxed{P8}$, one addressing the accuracy-privacy trade-off of differentially private recommender systems $\boxed{P9}$, and one surveying the use of *differential privacy* in collaborative filtering recommender systems $\boxed{P10}$.

In the third field, Table 1.1 contains two publications that study popularity bias and characteristics of "niche" users in music recommendations $\boxed{P11}$ $\boxed{P12}$. One paper further studies if users of different genders are equally affected by popularity bias in music recommendations $\boxed{P13}$, and another paper studies popularity bias in multimedia recommendation domains $\boxed{P14}$. Furthermore, this list contains an online study on popularity bias mitigation in news article recommendations $\boxed{P15}$. Another paper analyzes miscalibration and popularity bias amplification in recommendations $\boxed{P16}$. Finally, one journal article studies long-term dynamics of fairness in algorithmic decision support in the labor market $\boxed{P17}$.

Table 1.1: List of main publications selected by the author of this habilitation.

No.	Publication
	Transparency and Cognitive Models in Recommender Systems
$\boxed{P1}$	Seitlinger, P., Ley, T., **Kowald, D.**, Theiler, D., Hasani-Mavriqi, I., Dennerlein, S., Lex, E., Albert, D. (2018). Balancing the Fluency-Consistency Tradeoff in Collaborative Information Search with a Recommender Approach. *International Journal of Human–Computer Interaction*, 34:6, pp. 557-575. DOI: https://doi.org/10.1080/10447318.2017.1379240

$P2$	Lex, E.*, **Kowald, D.***, Schedl, M. (2020). Modeling Popularity and Temporal Drift of Music Genre Preferences. *Transactions of the International Society for Music Information Retrieval*, 3:1, pp. 17-30. (*equal contribution) DOI: https://doi.org/10.5334/tismir.39
$P3$	**Kowald, D.***, Lex, E.*, Schedl, M. (2020). Utilizing Human Memory Processes to Model Genre Preferences for Personalized Music Recommendations. In *4th Workshop on Transparency and Explainability in Adaptive Systems through User Modeling Grounded in Psychological Theory (HUMANIZE @ ACM IUI'2020)*. (*equal contribution) DOI: https://doi.org/10.48550/arXiv.2003.10699
$P4$	Lex, E., **Kowald, D.**, Seitlinger, P., Tran, T., Felfernig, A., Schedl, M. (2021). Psychology-informed Recommender Systems. *Foundations and Trends in Information Retrieval*, 15:2, pp. 134–242. DOI: https://doi.org/10.1561/1500000090
$P5$	Moscati, M., Wallmann, C., Reiter-Haas, M., **Kowald, D.**, Lex, E., Schedl, M. (2023). Integrating the ACT-R Framework and Collaborative Filtering for Explainable Sequential Music Recommendation. In *Proceedings of the 17th ACM Conference on Recommender Systems (RecSys'2023)*, pp. 840–847. DOI: https://doi.org/10.1145/3604915.3608838

Privacy and Limited Preference Information in Recommender Systems

$P6$	Duricic, T., Lacic, E., **Kowald, D.**, Lex, E. (2018). Trust-Based Collaborative Filtering: Tackling the Cold Start Problem Using Regular Equivalence. In *Proceedings of the 12th ACM Conference on Recommender Systems (RecSys'2018)*, pp. 446–450. DOI: https://doi.org/10.1145/3240323.3240404
$P7$	Lacic, E., Reiter-Haas, M., **Kowald, D.**, Dareddy, M., Cho, J., Lex, E. (2020). Using Autoencoders for Session-based Job Recommendations. *User Modeling and User-Adapted Interaction*, 30, pp. 617–658. DOI: https://doi.org/10.1007/s11257-020-09269-1
$P8$	Muellner, P., **Kowald, D.**, Lex, E. (2021). Robustness of Meta Matrix Factorization Against Strict Privacy Constraints. In *Proceedings of the 43rd European Conference on Information Retrieval (ECIR'2021)*, pp. 107-119. DOI: https://doi.org/10.1007/978-3-030-72240-1_8
$P9$	Muellner P., Lex, E., Schedl, M., **Kowald, D.** (2023). ReuseKNN: Neighborhood Reuse for Differentially-Private KNN-Based Recommendations. *ACM Transactions on Intelligent Systems and Technology*, 14:5, pp. 1-29. DOI: https://doi.org/10.1145/3608481
$P10$	Muellner P., Lex, E., Schedl, M., **Kowald, D.** (2023). Differential Privacy in Collaborative Filtering Recommender Systems: A Review. *Frontiers in Big Data*, 6:1249997, pp. 1-7. DOI: https://doi.org/10.3389/fdata.2023.1249997

	Fairness and Popularity Bias in Recommender Systems
P11	**Kowald, D.**, Schedl, M., Lex, E. (2020). The Unfairness of Popularity Bias in Music Recommendation: A Reproducibility Study. In *Proceedings of the 42nd European Conference on Information Retrieval (ECIR'2020)*, pp. 35-42. DOI: https://doi.org/10.1007/978-3-030-45442-5_5
P12	**Kowald, D.**, Muellner, P., Zangerle, E., Bauer, C., Schedl, M., Lex, E. (2021). Support the Underground: Characteristics of Beyond-Mainstream Music Listeners. *EPJ Data Science*, 10:14. DOI: https://doi.org/10.1140/epjds/s13688-021-00268-9
P13	Lesota, O., Melchiorre, A., Rekabsaz, N., Brandl, S., **Kowald, D.**, Lex, E., Schedl, M. (2021). Analyzing Item Popularity Bias of Music Recommender Systems: Are Different Genders Equally Affected? In *Proceedings of the 15th ACM Conference on Recommender Systems (RecSys'2021)*, pp. 601-606. DOI: https://doi.org/10.1145/3460231.3478843
P14	**Kowald, D.**, Lacic, E. (2022). Popularity Bias in Collaborative Filtering-Based Multimedia Recommender Systems. In *Advances in Bias and Fairness in Information Retrieval (BIAS @ ECIR'2022)*. Communications in Computer and Information Science, vol. 1610, pp. 1-11. DOI: https://doi.org/10.1007/978-3-031-09316-6_1
P15	Lacic, E., Fadljevic, L., Weissenboeck, F., Lindstaedt, S., **Kowald, D.** (2022). What Drives Readership? An Online Study on User Interface Types and Popularity Bias Mitigation in News Article Recommendations. In *Proceedings of the 44th European Conference on Information Retrieval (ECIR'2022)*, pp. 172-179. DOI: https://doi.org/10.1007/978-3-030-99739-7_20
P16	**Kowald, D.***, Mayr, G.*, Schedl, M., Lex, E. (2023). A Study on Accuracy, Miscalibration, and Popularity Bias in Recommendations. In *Advances in Bias and Fairness in Information Retrieval (BIAS @ ECIR'2023)*. Communications in Computer and Information Science, vol. 1840, pp. 1-16. (*equal contribution) DOI: https://doi.org/10.1007/978-3-031-37249-0_1
P17	Scher, S., Kopeinik, S., Truegler, A., **Kowald, D.** (2023). Long-Term Dynamics of Fairness: Understanding the Impact of Data-Driven Targeted Help on Job Seekers. *Nature Scientific Reports*, 13:1727. DOI: https://doi.org/10.1038/s41598-023-28874-9

I have contributed substantially to all 17 publications, and for 10 of these publications I am also either first or last author. The full texts of these publications can be found when following the respective DOI links in Table 1.1. For a full list of my publications, please take a look at my Google Scholar profile: https://scholar.google.at/citations?user=qQ-L8rUAAAAJ&hl=en.

Chapter 2

Related Work and Background

This chapter describes relevant research and background related to the scientific contributions of this habilitation. First, the main concepts of recommender systems are briefly outlined in Section 2.1, followed by relevant background with respect to transparency and cognitive models in recommender systems in Section 2.2. Next, the topic of privacy and limited preference information in recommender systems is briefly discussed in Section 2.3. Finally, Section 2.4 gives a compact overview of fairness and popularity bias in recommender systems. This chapter also summarizes our own research related to these topics, which is then outlined in relation to the main publications of this habilitation in Chapter 3.

2.1 Main Concepts of Recommender Systems

This section gives a compact overview of recommender systems (i) algorithms, (ii) applications, and (iii) evaluation methods relevant to this habilitation.

Recommender Systems Algorithms

In general, there are three main categories of recommendation algorithms [18,231]: (i) collaborative filtering (CF), (ii) content-based filtering (CBF), and (iii) hybrid approaches. This habilitation focuses on CF, but also investigates CBF.

Typically, a user-based CF recommender system $\mathcal{R}^k$ generates an estimated rating score for a target user u and a target item i by utilizing the ratings $r_{n,i}$ of k other users that have rated i, i.e., the k nearest neighbors $N_{u,i}^k$ [68]. Therefore, this variant of CF is often referred to as *UserKNN*, i.e., user-based k nearest neighbors. Formally, the estimated rating score $\mathcal{R}^k(u,i)$ for u and i is given by:

$$\mathcal{R}^k(u,i) = \frac{\sum_{n \in N_{u,i}^k} sim(u,n) \cdot r_{n,i}}{\sum_{n \in N_{u,i}^k} sim(u,n)} \tag{2.1}$$

where $sim(u,n)$ is the similarity between target user u and neighbor n. For *UserKNN*, the neighborhood $N_{u,i}^k$ used for generating recommendations for u and i comprises the k most similar neighbors. More formally:

$$N_{u,i}^k = \arg\max_{c \in U_i}^{k} sim(u,c) \tag{2.2}$$

where U_i are all users that have rated i and *sim* is the similarity metric (e.g., Cosine or Pearson [35]). There also exist variations of this algorithm suitable for item relevance prediction and for implicit user preferences (e.g., clicks) [123].

It is also possible to calculate similarities between items based on users' preferences of these items. This variant of CF is termed item-based CF (or *ItemKNN*) and has advantages in cases when user profiles change quickly [234]. *ItemKNN* was introduced as the main recommendation algorithm by Amazon.com [179] in 2003. For a comprehensive review of KNN-based CF methods, please see [210], and for a survey on CF with side information, please see [249]. Another possibility to incorporate side and context information (e.g., time or location) is by utilizing context-aware recommender systems, as discussed in these works [14, 17, 19].

Another variant of CF is *matrix factorization* (MF), which follows the idea that a user's preferences can be efficiently represented in low-dimensional space [140, 248]. The items are represented in the same low-dimensional space, which enables to generate recommendations by calculating the dot-product between the user and the item vectors. These vectors are often termed *embeddings*, and can be calculated with techniques such as graph neural networks [258], recurrent neural networks [113], neural CF [109], or autoencoders [284]. As described in Section 2.3, in this habilitation, autoencoders are used to address the issue of limited preference information in session-based recommendations [124, 175]. Furthermore, a neural CF approach [109] is used to study differentially-private recommendations. For a comparison of neural network and KNN-based methods, please see [88].

The next type of algorithms, content-based filtering (CBF) [181], utilizes content features of items (e.g., genres, title) to build item profiles to overcome the item cold-start problem (i.e., items with no user preference information). These item profiles are then matched with user profiles that also consist of content features of the consumed items [63]. For representing content features, techniques such as LDA (Latent Dirichlet Allocation) [43] can be used. CBF could suffer from a lack of novelty and diversity, since typically items are recommended that are similar to the items the user has consumed in the past. To overcome this issue, hybrid recommendation approaches [45, 46, 141, 219] combine CBF and CF to get the benefits of both worlds. There exist several ways to combine recommendation algorithms [128], including (i) monolithic, where collaborative and content information is combined in a single recommendation model, (ii) parallelized, where the results of different algorithms are combined using, e.g., a weighted approach, and (iii) pipelined, where one algorithm uses the results of another algorithm as input.

Recommender Systems Applications

This habilitation focuses on four application areas for recommender systems, namely (i) tag recommendations, (ii) music recommendations, (iii) job recommendations, and (iv) news article recommendations. The following paragraphs briefly describe the particularities of these application areas. Tag recommendation systems aim to support users in finding descriptive tags (or keywords) for annotating Web resources [129, 185] (e.g., music tracks or tweets in Twitter). Previous research of this habilitation's author has shown that a user's choice of tags is affected by activation processes in human memory [149, 150], which can be utilized

for a transparent design of tag recommendation models [155] (see Section 2.2).

Similar to recommendations in other multimedia domains [66] (e.g., movies or television items [187]), music recommender systems help users to navigate large content databases, and to find content that suit their taste [240]. However, in contrast to movies or books, music has some distinguished properties that also affect the design of music recommendation algorithms [239]: (i) music may be consumed repeatedly, while movies or books are typically consumed only once or a few times at maximum, (ii) music recommendations can be addressed on different abstraction levels including tracks, albums, artists and genres, (iii) rating data is relatively rare in the music domain, and thus implicit user preferences (e.g., listening events) are an important information source for recommender systems [70], and (iv) domain knowledge (e.g., musical sophistication) may have a high impact on how recommendations are perceived by the music listeners [131].

Next, job recommender systems address a particular recommendation problem, in which open job positions should be matched with job candidates [10, 11]. This differs to other recommendation application domains, since typically every open job position (i.e., the item) can be assigned to only one job candidate (i.e., the user), and vice versa [134]. Additionally, job portals (especially those that offer jobs to students and young talents) often provide the possibility to browse jobs anonymously [165, 227], which then turns the job recommendation problem into a session-based recommendation problem [124, 223]. Limited preference information and anonymous user sessions are also issues of news article recommender systems [64, 117, 199]. Via providing recommendations of currently relevant news articles that match session information (e.g., clicks) of the user, news portals aim to increase user engagement, and to turn anonymous readers into paying subscribers [5]. Finally, another particularity of news recommendations is the short lifetime of items, since many articles are only relevant for one day [215].

Recommender Systems Evaluation Methods

This habilitation considers both online and offline evaluation procedures of recommender systems. Both methods aim to compare the performance of two or more recommendation algorithms, but while online evaluation is performed in a live system, e.g., using A/B tests [95], offline evaluation is performed using collected preferences, typically in the form of training and test sets [53]. Another difference lies in the time when the user preference information is collected: whereas online evaluation collects user preferences after the recommendations are shown to the users, offline evaluation gathers user preferences (i.e., the ground truth data in the test set) before the recommendations are calculated [52, 283]. Online evaluation procedures then measure the actual performance using impact- or value-oriented measurements such as *Click-Through-Rate (CTR)* [122]. In contrast, offline evaluation procedures make use of relevance or performance metrics, which are often borrowed from the information retrieval research field [28, 233].

With respect to offline evaluation metrics, this habitation investigates both accuracy and beyond-accuracy metrics. To measure accuracy [61], error-based metrics for rating prediction such as the *Mean Absolute Error (MAE)* [279], and metrics for ranking quality such as *Precision (P), Recall (R), F1-score (F1), Mean*

Reciprocal Rank (MRR), and *Normalized Discounted Cumulative Gain (nDCG)*, have been proposed in the literature (e.g., [112]).

After decades of accuracy-driven recommender systems evaluation procedures, the research community has argued that being accurate is not the only important objective for a recommender system, and has proposed a set of beyond-accuracy metrics [16, 72, 96, 192]. Here, especially, the concepts of *novelty* and *diversity* are important [51]. Novelty describes the difference between the recommended items and a specific context, which could be the target user's item history or all users' item histories in the system [272]. The former, which is also referred to as personalized or user-based novelty, or unexpectedness [13], describes how different the recommendation list is from the items the target user has consumed in the past (i.e., the user item history). This concept is also related to serendipity, which, in addition, takes the relevance of the recommended items into account [56]. The latter, which is also referred to as long-tail novelty or system-based novelty [272], measures the rarity or inverse popularity of the recommended items [54]. This concept is also related to evaluating fairness and popularity bias of recommendations, which is described in more detail in Section 2.4.

All methods and metrics discussed so far solely evaluate the recommender system from a user's, or consumer's, perspective. However, in recent years, the multi-stakeholder nature of recommender systems has been highlighted, which not only takes the users, but also the item providers (and maybe other stakeholders like the system operators) into account [1, 3, 252, 253]. Here, especially integrating and evaluating item provider constraints is becoming an important research topic [259], and is also related to multi-sided fairness aspects of recommender systems [48, 254]. Finally, the reproducibility of recommender systems evaluation procedures is another important and timely topic [60]. Here, the adequate documentation and sharing of source-code and dataset samples used in the evaluation process is a key aspect of reproducibility [32]. Please see Section 3.4, for a discussion of reproducibility aspects related to the contributions of this habilitation.

> **Summary of own research (1):** This habilitation studies a wide range of recommendation algorithms and applications such as tag, music, job, and news article recommendations. Additionally, we investigate both accuracy and beyond-accuracy metrics, and both offline and online evaluation settings. Finally, we discuss reproducibility aspects of the scientific contributions of this habilitation, and provide code and data resources to foster reproducibility.

2.2 Transparency and Cognitive Models in Recommender Systems

This habilitation investigates transparency aspects of recommender systems by following principles of psychology and human cognition for a transparent design process of recommendation algorithms. Another possibility to enhance transparency in recommender systems is by providing explanations for recommendations, which is not investigated in this habilitation. For the field of explainability in recommender systems, please see [213, 261, 262, 263].

2.2.1 The Role of Psychology in Recommender Systems

Already, early research in the field of recommender systems was influenced by the fact that humans' decision-making processes are impacted by their social surroundings, which also motivated the implementation of the first collaborative filtering-based recommendation algorithms [229, 230]. In order to create more human-centric recommendations, additional psychological characteristics of users were incorporated in the design and implementation process of recommender systems [27, 265]. For example, insights from decision psychology [170] were used to study serial position and anchoring effects in recommendations [15, 86, 256], and to show that users are more likely to remember items at the beginning (i.e, *primacy* effect) and the end (i.e., *recency* effect) of a list [257]. Related research also investigated how to incorporate aspects such as personality [264], and affect, e.g., emotion [100] or satisfaction [188, 191], into the recommendation process.

Based on these lines of research, we survey and categorize related work at the intersection of psychology and recommender systems. We term this type of recommender systems *psychology-informed recommender systems* [174], and we identify three main areas: (i) cognitive (or cognition)-inspired, (ii) personality-aware, and (iii) affect-aware recommender systems. Additionally, we connect these areas to aspects of human decision-making, and to aspects of human-centric evaluation design of recommender systems. This habilitation focuses on the first area, namely cognitive-inspired recommendations based on human memory theory, which is described in more detail in the following section.

> **Summary of own research (2):** We highlight the usefulness of incorporating the underlying psychological constructs and theories into a transparent design process of recommender systems. We term this type of recommender system *psychology-informed recommender system*, and categorize it into three types.

2.2.2 Cognitive-inspired Recommendations

This habilitation investigates two cognitive-inspired recommendation approaches: one based on human episodic memory, and another one based on activation processes in human memory. Other types of cognition-aware recommendation approaches, such as stereotype-based recommendations [232], categorization-based recommendations [245], or attention-based user models [244], are discussed in [174].

Recommendations based on Human Episodic Memory

Human episodic memory is the memory of personally experienced events that occurred in a specific context (e.g., a particular day or place, or a given categorization) [268]. The contextual information is essential for retrieving these events. MINERVA2 [114] is a model that accounts for episodic memory-based human behavior such as categorization [115], and recognition [116]. MINERVA2 distinguishes between a long-term or secondary memory that holds the episodic memory traces (i.e., the events along with the context information), and a working or primary memory that communicates with the secondary memory by sending retrieval cues (e.g., current context information), and receiving matching events.

In our own research [158, 246], we employ MINERVA2 to implement a tag recommendation algorithm called *Search of Memory (SoMe)*. *SoMe* mimics a user's search of memory when assigning tags to bookmark a Web resource. Therefore, we encode episodic memory traces using the categories assigned to previously bookmarked Web resources of this user. Specifically, *SoMe* implements MINERVA2's distinction between the primary and secondary memory in a way that the primary memory represents the Web resource to be tagged in terms of the resource's categories, and to search the secondary memory for tags that are assigned to Web resources with similar categories. These tags are then recommended to the user. Via user studies, we find that *SoMe* provides higher tag recommendation acceptance than a popularity-based baseline approach [158, 246] (see Section 3.1).

Recommendations based on Activation Processes in Human Memory

Human memory is very efficient in making memory units quickly available when they are needed [39, 217]. More formally, human memory tunes the activation of its units to statistical regularities of the current context and environment [24]. These so-called activation processes in human memory are formalized in the cognitive architecture ACT-R [23]. ACT-R is short for "Adaptive Control of Thought – Rational", and differs between two long-term memory modules: (i) declarative memory, which holds factual knowledge (i.e., what something is), and (ii) procedural memory, which consists of action sequences (i.e., how to do something) [22].

This habilitation focuses on the declarative memory module, which contains the *activation equation* of human memory. The *activation equation* determines the usefulness, i.e., the activation level A_i, of a memory unit i (e.g., a specific item or item category the user has interacted with in the past) for a user u in the current context. It combines a *base-level* activation with an *associative* activation, which depends on the weight W_j, and the strength of association $S_{j,i}$ [22]:

$$A_i \;=\; B_i + \sum_j W_j \cdot S_{j,i} \tag{2.3}$$

where B_i represents the *base-level* activation of i, which quantifies its general usefulness by considering how frequently and recently it has been used in the past. It is defined by the *base-level learning (BLL) equation* [24]:

$$B_i \;=\; ln\left(\sum_{j\,=\,1}^{n} t_j^{-d}\right) \tag{2.4}$$

where n is the frequency of i's occurrences in the past (i.e., how often u has interacted with i), and t_j is the time since the jth occurrence of i (i.e., the recency of i). The exponent d accounts for the time-dependent decay of item exposure, which means that each unit's activation level decreases in time according to a power function. The second part of Equation 2.3 represents the *associative activation* that tunes B_i to the current context. The current context can be defined by any contextual element j that is relevant to the current situation, and via learned associations, the contextual elements can increase i's activation.
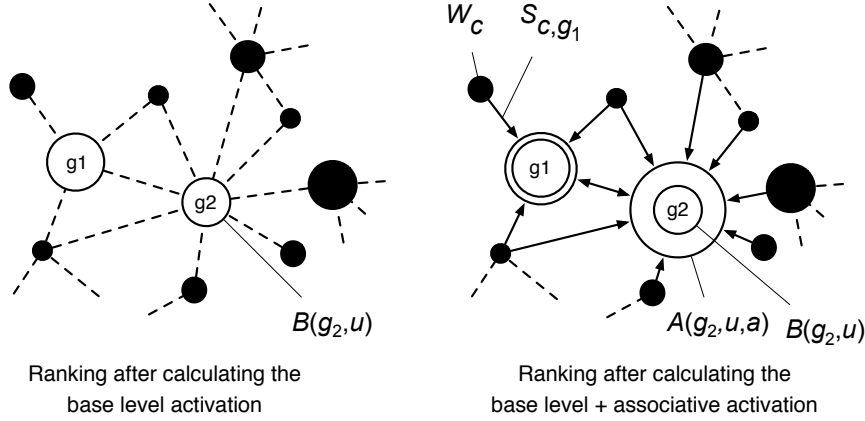
Figure 2.1: An example illustrating the difference between the *BLL equation* (left panel) and the *activation equation* (right panel). Here, unfilled nodes represent target genres g_1 and g_2, and black nodes represent contextual genres. For g_1 and g_2, the node sizes represent the activation levels, and for the contextual genres, the node sizes represent the weights W_c. The association strength $S_{c,g}$ is represented by the length of each edge. We see a different ranking of the genres in the two settings, which illustrates the importance of the associative activation [152, 267].

Figure 2.1 illustrates the difference between the base-level activation and the associative activation in the case of a music recommendation system that aims to rank relevant music genres for a given user. The left panel shows the ranking of two genres g_1 and g_2 according to the *BLL equation*. Here, g_1 would have a higher activation level than g_2 based on past usage frequency and recency. The right panel shows the ranking of these genres according to the *activation equation*, which also takes associations with contextual genres into account (e.g., music genres that are relevant in the current situation). Using the combined base-level and associate activation, the ranking changes, and g_2 would have a higher activation level than g_1 [152]. The declarative memory module also contains some additional components. One example is the valuation component [132], which determines the *value* attributed by u to i (e.g., interaction time or frequency [226]).

In our research, we use the *BLL equation* and *activation equation* for a transparent design, implementation, and evaluation process of two music recommendation algorithms [152, 173]. We show that these cognitive-inspired approaches outperform related baselines in terms of recommendation accuracy. Additionally, we illustrate to what extent the components of ACT-R contribute to the generation of the music recommendation lists [201]. In a recently accepted paper [156], we discuss transparency aspects of additional components of ACT-R.

Summary of own research (3): We use models of human episodic memory (i.e., MINERVA2), and activation process in human memory (i.e., ACT-R) for a transparent design, implementation, and evaluation process of recommendation algorithms. We also illustrate to what extent the components of ACT-R (e.g., *BLL*) contribute to the generation of the recommendation lists.

2.3 Privacy and Limited Preference Information in Recommender Systems

This section gives an overview of privacy-aware recommendations. Since the users' privacy concerns could also lead to the limited availability of preference information (e.g., users disclose their preferences, or do not sign in to the system), this section also gives an overview of session-based and cold-start recommendations.

2.3.1 Privacy-aware Recommendations

In terms of privacy, this habilitation focuses on differentially-private recommendations. This section also briefly discusses privacy aspects of recommender systems.

Privacy Aspects of Recommender Systems

Recommender systems need to store and process user preference information, which could lead to potential privacy risks to its users [91]. This includes the inference of private information. Here, related research has shown that inference attacks can be used to derive a user's sensitive information (e.g., gender [251]) based on the information shared with the recommender system [33, 130, 278]. For example, in k-nearest neighbor-based recommender systems, the use of neighbors' preference information in the recommendation process can pose a privacy risk to the neighbors [224, 286]. In this way, the preference information of the neighbors can be uncovered, or the neighbors' identities (or sensitive attributes) can be revealed. Other inference attacks in recommender systems work by generating fake users, i.e., sybils, based on the limited knowledge of a victim's preferences. These sybils isolate the victim utilized as a neighbor, and compromise its privacy [49].

Different privacy-preserving technologies have been used to mitigate the users' privacy risks, including *homomorphic encryption*, *federated learning*, and *differential privacy*. While *homomorphic encryption* techniques aim to generate privacy-aware recommendations by employing encrypted user preference information [285], *federated learning* techniques build on the assumption that sensitive user information should never leave the user's device [25, 282]. Finally, *differential privacy* protects the users by introducing noise into the recommendation process [74].

Our own research focuses on using *differential privacy*. Additionally, we study how limiting the preference information of users can help to increase privacy. Therefore, we use the concept of *meta learning* [178] to calculate recommendations based on a minimal amount of user preference information. With this, we study privacy constraints of users (e.g., willingness to share preference information) [202]. We find that users with small profiles can afford a higher degree of privacy than users with large profiles, and that *meta learning* is helpful for increasing the robustness against the users' privacy constraints (see Section 3.2).

Differentially-Private Recommendations

The aim of differentially-private recommendations is to inject randomness and noise into the recommendation calculation process to mitigate the inference risk

of users' preference information [90, 193]. This habilitation focuses on a specific attack, which can be addressed by using *differential privacy*. Here, a user with malicious intent, i.e., the *adversary* a, tries to infer preference information (here, rating scores) of a specific neighbor n in user-based k-nearest neighbor CF (i.e., *UserKNN*) [49]. In this attack scenario, the adversary a has some prior knowledge about n, such as publicly available rating information P of n that could have been inferred from, e.g., product reviews. Using P, a modifies its own user profile R_a such that it (partially) matches n's profile, which increases the likelihood of n being used as a neighbor for calculating a's recommendations. With this, a queries estimated rating scores from the recommender system, i.e., $\mathcal{R}^k(a) = \{\mathcal{R}^k(a, i_1), \mathcal{R}^k(a, i_2), \dots, \mathcal{R}^k(a, i_l)\}$, where $\mathcal{R}^k(a, i_j)$ is the estimated rating score for item $i_j \in Q_a$, and Q_a is the set of a's rating queries. Then a aims to infer rating information r_{n,i_j} of a neighbor n for item i_j used to generate the estimated rating scores. More formally, this is given by:

$$Pr[r_{n,i_1}, r_{n,i_2}, \dots, r_{n,i_l} | \mathcal{R}^k(a, i_1), \mathcal{R}^k(a, i_2), \dots, \mathcal{R}^k(a, i_l), P \cup R_a] \quad (2.5)$$

To mitigate the inference risk of n's rating information, different variants of *differential privacy* such as the *Laplace input perturbation* [76] or *plausible deniability* [41] can be used. This habilitation utilizes *randomized responses* [277] to establish *plausible deniability*. Specifically, a privacy mechanism m_{DP} is applied to the neighbors' ratings to generate the differentially-private set of ratings $\tilde{R}$:

$$\tilde{R} = \{m_{DP}(r_{n,i}) : n \in N_{u,i}^k\} \quad (2.6)$$

Via *randomized responses*, neighbors can plausibly deny that their real rating was used in the recommendation process. In detail, the privacy mechanism m_{DP} flips a fair coin, and if the coin is heads, the neighbor's real rating is used in the recommendation calculation. If the coin is tails, m_{DP} flips a second fair coin to decide whether the neighbor's real rating, or a random rating drawn from a uniform distribution over the range of ratings, is used. With this, the adversary a does not know if the utilized rating is real, or random, which leads to the guarantees of *differential privacy* [76]. However, the randomness introduced to the users' preference information typically leads to accuracy drops, and thus also to a fundamental trade-off between accuracy and privacy [38].

In our research, we address this accuracy-privacy trade-off by proposing a novel differentially-private recommendation approach termed *ReuseKNN* [206]. *ReuseKNN* aims to reduce the number of users that need to be protected via *differential privacy* by employing an efficient neighborhood reuse concept. With this, the majority of users (we call them *secure* users) are rarely used in the recommendation process and thus, do not need protection, while some highly reusable users (we call them *vulnerable* users) can be protected with *differential privacy*. Figure 2.2 schematically illustrates our approach, and shows that the fraction of *secure* users is substantially larger in the case of *ReuseKNN* compared to traditional *UserKNN*. We also find that this leads to higher recommendation accuracy compared to a fully differentially-private recommender system (see Section 3.2). Additionally, we survey, analyze, and categorize the use of *differential privacy* in 26 papers published in recommender systems-relevant venues [205].
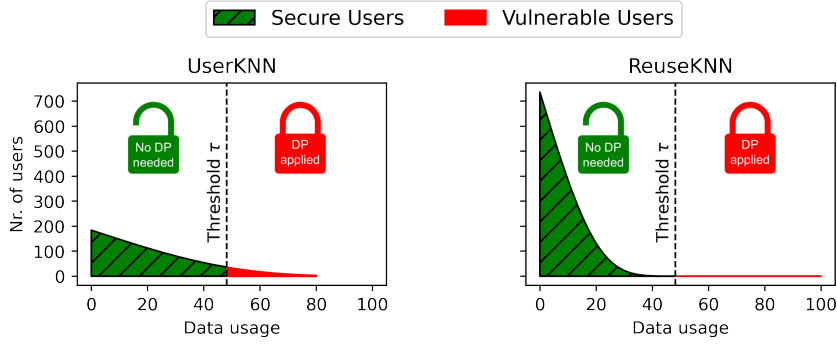
Figure 2.2: Schematic illustration of the data usage (i.e., how often a user is used as a neighbor) distribution of *UserKNN* and *ReuseKNN*. *ReuseKNN* increases the number of *secure* users (green, no *differential privacy* needed) and decreases the number of *vulnerable* users (red, *differential privacy* needs to be applied) compared to *UserKNN*. The dashed line illustrates the data usage threshold τ, a parameter to adjust the maximum data usage for users to be treated as secure.

> **Summary of own research (4):** We study varying privacy constraints of users, e.g., the willingness to share preferences with the recommender system. Additionally, we address the privacy-accuracy trade-off in differentially-private recommendations by employing a neighborhood reuse concept, and survey and categorize the literature on using *differential privacy* in collaborative filtering.

2.3.2 Limited Availability of User Preference Information

Data protection initiatives as well as the users' privacy concerns in recommender systems can lead to the limited availability of preference information [40,138,169, 194,270]. This habilitation investigates this issue in session-based and cold-start recommendation settings, which are discussed in this section.

Session-based Recommendations

Session-based recommender system aim to provide meaningful recommendations in cases where long-term user preferences, or user histories, are not available (e.g., due to users' privacy concerns, or when users do not sign in to the system). The input of a session-based recommender system consists of a typically short item sequence that is observed in the current user session [125,182]. Different algorithms for session-based recommendations have been proposed, including methods based on k-nearest neighbors [124] or recurrent neural networks [113]. Session-based recommender systems are related to sequence-aware and sequential recommender systems [133], which are not covered in this habilitation. Please see [222] for a detailed overview of sequence-aware and sequential recommendations.

In our research, we employ autoencoders, a specific type of neural network for reducing the dimensionality of data [160], to infer latent session representations, and to generate session-based job recommendations. Specifically, we find that variational autoencoders provide the best results across a set of accuracy and beyond-accuracy evaluation metrics (e.g., system and session-based novelty) [167].

The User Cold-Start Problem

The user cold-start problem in recommender systems refers to users that have interacted with only a few or even with no items at all, i.e., users with limited availability of preference information [241]. Related research has proposed different methods to address the user cold-start problem, including simple popularity-based and unpersonalized approaches [241], location-aware recommendations [166], and trust-based recommendations [78, 186]. This habilitation focuses on trust-based recommendations, which exploit trust statements between users (e.g., user A trusts user B) to create trust networks, and to calculate CF-based recommendations using the connections in these trust networks [104, 105, 214].

In our research, we employ network measures such as *regular equivalence* [110] to calculate trust-based recommendations for cold-start users. Via *regular equivalence*, we do not only find neighbors that share the same trust connections, but also neighbors that have similar structural roles in the trust network (e.g., users that are only connected to influential nodes in the network). We find that our approach outperforms related methods based on, e.g., Jaccard similarity [73].

> **Summary of own research (5):** We address the issue of limited availability of user preference information (e.g., due to users' privacy constraints) in session-based and cold-start recommendation settings. We demonstrate the usefulness of variational autoencoders for session-based job recommender systems. Furthermore, we address the user cold-start problem by employing trust-based recommendations using network measures such as *regular equivalence*.

2.4 Fairness and Popularity Bias in Recommender System

This section gives a brief overview of fairness in algorithmic decision support, and outlines research on popularity bias in recommender systems. For more detailed reviews on fairness-aware recommender systems, please see [65, 79, 276, 281].

2.4.1 Fairness in Algorithmic Decision Support

Fairness in algorithmic decision support and in machine learning applications has gained a lot of attention in recent years, and has been studied especially for binary classification problems [37, 161, 195]. In this problem setting, Y denotes the real outcome to be predicted by the classifier (e.g., the class label, for example if a job applicant has been put into a high- or low-prospect group), and A is the set of protected attributes of an individual, thus the attributes that one must not discriminate against (e.g., gender or race). Furthermore, X denotes non-protected attributes of an individual, and $\hat{Y}$ is the predictor of Y (e.g., to predict to which class the individual belongs), which could depend on X and A. Different definitions of fairness were proposed for such a setting in the literature.

For example, *fairness through unawareness* is satisfied if the predictor $\hat{Y}$ only depends on X and not on A to predict Y, i.e., $\hat{Y} : X \to Y$. Although this fairness

definition seems to be compelling and simple to implement, it was shown that it is not sufficient in the area of algorithmic decision support since elements of X may contain hidden discriminatory information of A (e.g., race may correlate with the place of residence) [106]. Another definition is based on *individual fairness* [75]. Given that we have a distance metric $d(i, j)$, if two individuals i (with X_i and A_i) and j (with X_j and A_j) are similar according to this metric (so $d(i, j)$ is small), then also their predicted outcomes should be similar: $\hat{Y}(X_i, A_i) \approx \hat{Y}(X_y, A_y)$. One drawback of *individual fairness* is that the definition of $d(i, j)$ requires detailed information of the individuals as well as detailed domain knowledge.

Apart from that, the literature has also provided different definitions for *group fairness*. According to the *statistical parity* (or demographic parity) definition [30], fairness is given if the positive outcome proportion of the predictor $P(\hat{Y} = 1)$ is equal for all A, which, in the binary case with $A \in \{0, 1\}$, is given by:

$$P(\hat{Y} = 1|A = 0) = P(\hat{Y} = 1|A = 1) \tag{2.7}$$

Legally, this metric is often related to the 4/5th rule [101]. This means that the positive outcome ratio between the protected group (i.e., $A = 0$) and the privileged group (i.e., $A = 1$) should be at least 0.8. For example, if the privileged group has a positive outcome proportion of 50%, then the protected group should have a positive outcome proportion of at least 40%. The downside of this metric is that it does not depend on the real outcome Y (only on the predictions $\hat{Y}$).

In contrast, *equality of opportunity* [107] also takes the real outcome Y into account. The idea is that individuals of the privileged and individuals of the protected group should have equal chance of getting a positive outcome, assuming that the individuals of the groups are qualified for this positive outcome. This can be measured via the true positive rate, which is given by:

$$P(\hat{Y} = 1|Y = 1, A = 0) = P(\hat{Y} = 1|Y = 1, A = 1) \tag{2.8}$$

Equality of opportunity can also be defined using the false negative rate [243]. Additionally, *equalized odds* is a stricter variant of *equality of opportunity* that requires that both the true positive rate and the false positive rate are equal [273]. Research has also found a trade-off between *individual* and *group fairness* [42].

In our research, we employ some of these definitions and adjust them to study long-term dynamics of fairness in algorithmic decision support. Therefore, we develop an agent-based model and evaluate it in a labor market setting [243]. We find that there is a trade-off between different long-term fairness goals, which validates the aforementioned *individual* and *group fairness* trade-off (see Section 3.3). Although, this work does not directly study recommender systems, it sheds light on the usefulness of agent-based modeling for studying algorithmic fairness in the long-term, which is also relevant for the research field of recommender systems.

> **Summary of own research (6):** We study long-term fairness dynamics in algorithmic decision support in a labor market setting using agent-based modeling techniques. We highlight the trade-off between different long-term fairness goals in such a setting (i.e., *individual* and *group fairness*).

2.4.2 Measuring, Understanding, and Mitigating Popularity Bias

In this section, metrics to measure and understand popularity bias, and methods to mitigate popularity bias in recommender systems are briefly discussed.

Popularity Bias Metrics

Research has shown that recommendation algorithms (especially those based on CF) are biased towards popularity, which leads to the overrepresentation of popular items in the recommendation lists [82,83]. This also leads to the underrepresentation of unpopular items (long-tail items) in the recommendation lists [44,218]. The literature has proposed different metrics to measure and understand popularity bias from the item and user perspective [20,137]. This habilitation focuses on three specific ways to measure inconsistencies between user groups with respect to popularity bias: (i) accuracy differences between user groups, (ii) *miscalibration*, and (iii) *popularity lift*. While the first one simply requires comparing the average recommendation accuracy between the groups, *miscalibration* and *popularity lift* are more complex to calculate. Additionally, via *skewness* and *kurtosis*, we measure the asymmetry and "tailedness" of the popularity distributions [34].

In general, *calibration* quantifies the similarity of a genre spectrum between a user profile p and a list of recommendations q [255]. For example, if a user has consumed 80% of rock music and 20% of pop music in the past, then a *calibrated* recommendation list should also contain this genre distribution. Although this metric is not a popularity bias metric by definition, it is often used to measure and understand popularity bias in recommendations [6,8]. The definition of *calibration* can be reinterpreted in the form of *miscalibration*, i.e., the deviation between p and q [177]. This deviation is calculated using the *Kullback-Leibler (KL)* divergence between the distribution of genres in p, i.e., $p(c|u)$, and the distribution of genres in q, i.e., $q(c|u)$. More formally, for user u, this is given by:

$$KL(p||q) = \sum_{c \in C} p(c|u) \log \frac{p(c|u)}{q(c|u)} \tag{2.9}$$

Here, C is the set of all genres in a given dataset. Therefore, $KL(p||q) = 0$ means perfect *calibration*, and higher $KL(p||q)$ values (i.e., close to 1) mean *miscalibrated* recommendations. The $KL(p||q)$ values can be averaged for a given group g.

In contrast, *popularity lift* measures to what extent recommendation algorithms amplify the popularity bias inherent in the user profiles [7,8]. Thus, this metric quantifies the disproportionate recommendation of popular items for a given user group g. *Popularity lift* is based on the group average popularity $GAP_p(g)$, which is defined as the average popularity of the items in the user profiles p of group g. Similarly, $GAP_q(g)$ is the average popularity of the recommended items for all users of the group g. Taken together:

$$PL(g) = \frac{GAP_q(g) - GAP_p(g)}{GAP_p(g)} \tag{2.10}$$

$PL(g) > 0$ means that g's recommendations are too popular, $PL(g) < 0$ means that g's recommendations are too unpopular, and $PL(g) = 0$ is the ideal value.

In our research, we use these metrics to study popularity bias in recommender systems [147,153,157]. We find that "niche" users interested in unpopular content receive worse recommendation quality than users interested in popular content. We study the characteristics of these "niche" users in the field of music recommendations, and identify subgroups that also differ in the recommendation quality they receive [154]. Finally, we also find that music recommendation algorithms could intensify the popularity bias for the group of female users [171].

Popularity Bias Mitigation

Research has proposed different methods to mitigate bias in algorithms, including pre-, in-, and post-processing methods [212]. In the field of recommender systems, especially in-processing and post-processing techniques are used to mitigate popularity bias. Here, in-processing algorithms aim to adjust the recommendation calculation procedure, and to correct the popularity bias using, e.g., *calibration*-based techniques [9, 136]. In contrast, post-processing techniques do not change the recommendation algorithm itself, but the generated recommendation list by using, e.g., re-ranking techniques [4, 26]. Typically, in-processing techniques are the most complex ones to implement, since the underlying algorithm needs to be adapted. However, they are efficient with respect to computational costs. In contrast, one drawback of post-processing techniques is the computational inefficiency of these methods due to the high computational complexity of item re-ranking. However, they can be applied to any given item ranking independent of the underlying algorithm [55]. Finally, the use of content-based recommendation algorithms [63, 181] is another possibility to address popularity bias in recommender systems due to their independence of user preference information [2, 203].

In our research, we study popularity bias mitigation in news article recommender systems for both subscribed users and anonymous session users utilizing content-based recommendations [163]. In an online study that we have conducted together with the Austrian news platform *DiePresse*, we find that personalized and content-based recommendations lead to a more balanced news article readership distribution compared to purely popularity-based recommendations. Thus, we find that readers are not only interested in the most popular and recent news articles, but also in long-tail articles if they match the user preference history, or the preferences tracked in the current session (see Section 3.3).

> **Summary of own research (7):** We analyze popularity bias in collaborative filtering-based recommender systems, and find that "niche" users interested in unpopular content receive worse recommendation accuracy than users interested in popular content. Thus, this "niche" user group is treated in an unfair way by collaborative filtering-based recommender systems. Furthermore, we analyze the characteristics of these users, and study popularity bias mitigation in news article recommender systems using content-based recommendations.

Please note that the aim of this "Related Work and Background" chapter has not been to give a comprehensive review of the various research fields mentioned, but rather to discuss the research and background related to the scientific contributions and publications of this habilitation described in the next chapter.

Chapter 3

Scientific Contributions

This chapter describes the scientific contributions of this habilitation according to the three research topics that are investigated: (i) transparency and cognitive models (Section 3.1), (ii) privacy and limited preference information (Section 3.2), and (iii) fairness and popularity bias (Section 3.3) in recommender systems. Therefore, the 17 publications listed in Table 1.1 are categorized into 7 scientific contributions. For research topics (i) and (ii), this leads to two contributions each, and for research topic (iii), this leads to three contributions, since this topic also covers the most publications of this habilitation. Additionally, Section 3.4 summarizes the scientific contributions, and elaborates on reproducibility aspects.

3.1 Transparency and Cognitive Models in Recommender Systems

This section summarizes our research on transparency aspects of recommendations by using cognitive modeling techniques. It contains three studies employing cognitive models for a transparent design process of tag and music recommendation algorithms $\boxed{P1}$ $\boxed{P2}$ $\boxed{P3}$, and one survey and categorization of psychology-informed recommender systems $\boxed{P4}$ (*Contribution 1*). Additionally, one study illustrates to what extent the components of the cognitive model ACT-R contribute to the generation of music recommendation lists $\boxed{P5}$ (*Contribution 2*).

Contribution 1: Using Cognitive Models for a Transparent Design and Implementation Process of Recommender Systems (2018-2021)

$\boxed{P1}$ introduces a tag recommendation algorithm termed *SoMe (Search of Memory)* based on *MINERVA2* [114], which is a model of human episodic memory (see Section 2.2.2). We implement *SoMe* using our *TagRec* framework [146, 148], and evaluate it in an online study with 18 participants. During the four-weeks study, the participants had to investigate a specific topic (i.e., "designing workplaces that inspire people") by collecting and tagging three topic-related Web resources per week. For this, the participants were supported with a social bookmarking user interface (based on the *KnowBrain* tool [67]) that contained support via tag rec-
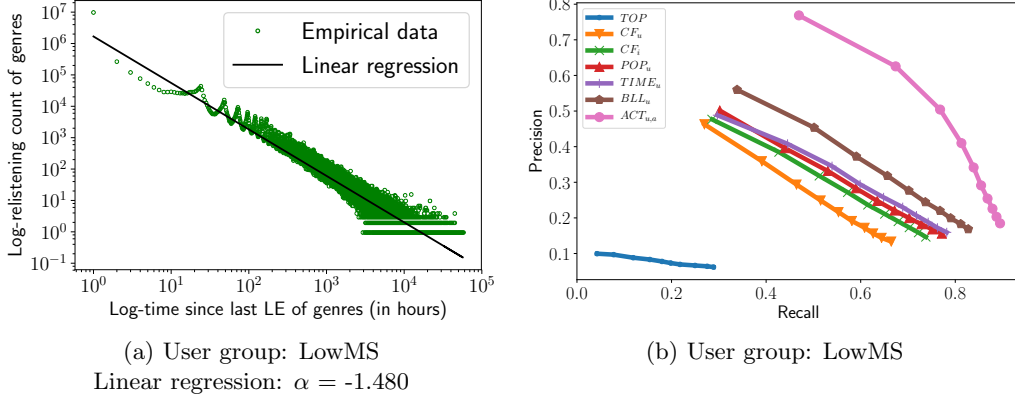
<div align="center">

(a) User group: LowMS
Linear regression: $\alpha = -1.480$
    (b) User group: LowMS

</div>

Figure 3.1: (a) Calculation of the BLL equation's d parameter. On a log-log scale, we plot the relistening count of the genres over the time since their last listening event (LE), and set d to the slopes α of the linear regression lines [173]. (b) Recall/precision plots for $k = 1 \dots 10$ predicted genres of the baselines, and our BLL_u and $ACT_{u,a}$ approaches. $ACT_{u,a}$ achieves the highest accuracy [152].

ommendations. Here, the participants randomly received tag recommendations calculated via *SoMe* or via a conventional *MostPopular* tag recommendation algorithm. Additionally, the participants were divided into two groups at random: (i) *individual*, where the participants only saw their own resources and tags, and (ii) *collaborative*, where the participants also saw the resources and tags of the other users in the group. Thus, in the *collaborative* setting, the tag recommendations were calculated based on the categorized resources and tags of the other users as well. The outcomes of our online study show that, in the *collaborative* setting, *SoMe* provides significantly higher tag recommendation acceptance rates than the *MostPopular* approach. In the *individual* setting, we do not observe a significant difference between the two approaches in terms of recommendation acceptance. Therefore, we find that a cognitive-inspired tag recommendation algorithm based on a transparent model of human episodic memory supports users in *collaborative* tagging settings. We have validated these findings in a follow-up paper using a similar tag recommendation approach termed *3Layers*, which we have presented at the *International World Wide Web conference 2018 (TheWebConf)* [158].

P2 and P3 present the second set of our cognitive-inspired recommendation algorithms based on activation processes in human memory as defined by the cognitive architecture ACT-R [22] (see Section 2.2.2). We introduce two algorithms for a transparent modeling and prediction approach for music genre preferences of users: (i) BLL_u, which implements the *base-level learning* (BLL) equation of ACT-R as described in P2, and (ii) $ACT_{u,a}$, which extends BLL_u, and implements the full *activation equation* of ACT-R as described in P3. We evaluate these approaches using dataset samples containing preferences (listening events) of users of the Last.fm music platform, based on the *LFM-1b* dataset [235, 238].

Figure 3.1 (a) illustrates the impact of time on the re-listening behavior of users in our Last.fm dataset sample. We find that users tend to listen to music

genres to which they have listened to very recently, and that this temporal decay follows a power-law distribution as suggested by the BLL equation of ACT-R [24]. We use the slope α of the linear regression of this data to set BLL's d parameter. Figure 3.1 (b) shows the accuracy of our approaches compared to five baseline algorithms: TOP suggests the most popular genres in the system, CF_u and CF_i represent user-based and item-based CF, and POP_u and $TIME_u$ suggest the most popular and most recent genres listened to by u, respectively. We find that BLL_u outperforms all baselines, and that $ACT_{u,a}$ outperforms BLL_u by also taking into account the current context of music listening (i.e., the genres of the artist a to which the user u listened to most recently) via the spreading activation component. Our findings show the usefulness of activation processes in human memory for a transparent design process of music recommendation algorithms, which also leads to high recommendation accuracy. We have validated these findings for the task of hashtag recommendations [155, 172], and for the task of music artist recommendations, which we have presented at the *International Society for Music Information Retrieval (ISMIR)* conference 2019 [151].

Finally, $\boxed{P4}$ surveys and categorizes recommender systems that draw on psychological theories for a transparent design, implementation, or evaluation process of recommendations. We term this type of recommender systems *psychology-informed recommender systems* and categorize them into three groups: (i) cognition (or cognitive)-inspired, (ii) personality-aware, and (iii) affect-aware recommender systems (see Section 2.2.1). We also discuss open issues in this research field, for example, the need to incorporate psychological considerations into the design process of user-centric recommender system evaluation studies.

Contribution 2: Illustrating to What Extent Components of the Cognitive Model ACT-R Contribute to Recommendations (2022-2023)

In $\boxed{P5}$, we discuss transparency aspects of music recommendations generated via ACT-R by illustrating to what extent components of ACT-R have contributed to the generation of recommendation lists. We investigate three ACT-R components described in Section 2.2.2: (i) *the base level learning* (BLL) equation, which describes the "current obsession" of a user (i.e., frequently and recently listened tracks), (ii) the spreading activation (S) component, which describes "current vibes" of a user (i.e., tracks that are similar to the user's most recently listened track), and (iii) the valuation (V) component, which accounts for "evergreens" of the user (i.e., the user's most frequently listened tracks, independent of the recency component). Additionally, we analyze a social component (SC) to account for track recommendations "from similar listeners" in the form of user-based CF.

Figure 3.2 shows six recommended tracks for a randomly selected user in our newly created Last.fm dataset sample [200] based on the *LFM-2b* dataset [196, 237]. The heatmap illustrates how the music track recommendations are calculated by showing the relative contribution of these four components to the recommendation score of a track. We see that the components contribute differently to the recommended tracks. For example, for the first track "From the Past Comes the Storms", the current obsession (BLL) of the user is most important, while for the last track "Troops of Doom" solely the social component (SC) contributes to

Recommended Track	Current obsession (BLL)	Current vibes (S)	Evergreens (V)	From similar listeners (SC)
From the Past Comes the Storms	0.471	0.248	0.281	0.000
Escape to the Void	0.306	0.353	0.341	0.000
To the Wall	0.294	0.359	0.347	0.000
R.I.P. (Rest in Pain)	0.264	0.374	0.362	0.000
The Abyss	0.263	0.375	0.362	0.000
Troops of Doom	0.000	0.000	0.000	1.000

Figure 3.2: Heatmap illustrating the relative contribution of three ACT-R components (BLL, S, and V) and one social component (SC) to the recommendation scores of six recommended tracks for a randomly chosen Last.fm user [201].

the recommendation calculation. Based on this, concrete explanations could be derived for all recommendations generated with this model. For example, "this track was recommended because of your *current obsession*", or "this track was recommended because of *similar listeners*". We discuss transparency aspects of additional components of ACT-R for music recommendations in a chapter for the "*A Human-centered Perspective of Intelligent Personalized Environments and Systems*" Springer book, which was recently accepted for publication [156].

3.2 Privacy and Limited Preference Information in Recommender Systems

Limited availability of user preference information (e.g., clicks) could be one consequence of data protection initiatives or of the users' privacy concerns in recommender systems [40, 138, 169, 194, 270] (e.g., users are not willing to share preferences, or to sign in to the system). Thus, we discuss the findings of two studies that address the limited availability of user preference information in the settings of session-based and cold-start recommendations P6 P7 (*Contribution 3*). Additionally, we address varying privacy constraints of users in recommender systems (e.g., hiding preferences) P8, and the accuracy-privacy trade-off of differentially-private recommender systems P9. Finally, we survey and categorize the literature on *differential privacy* in collaborative filtering P10 (*Contribution 4*).

Contribution 3: Addressing Limited User Preference Information in Cold-Start and Session-based Recommendation Settings (2018-2020)

P6 presents a trust-based CF approach for addressing the user cold-start problem in recommender systems (see Section 2.3.2). Specifically, we aim to exploit implicit and explicit connections between users in trust networks [186] to find the k nearest neighbors and to overcome the limited availability of user preference information in this setting. By employing the idea of *regular equivalence* via *Katz* similarity [110], we do not only find neighbors that share the *same* trust connections, but also neighbors that have *similar* trust connections (i.e., neighbors with similar structural roles in the network). We evaluate our approach using a dataset

from the consumer reviewing portal Epinions [186], which allows users to specify trust connections to other users. We find that our approach outperforms related approaches (e.g., based on Jaccard similarity [58]) in terms of recommendation accuracy for cold-start users. In our follow-up work [71], we employ graph embedding techniques on the trust network of users by evaluating graph embedding methods such as *graph factorization* [21], *DeepWalk* [220], or *Node2Vec* [102] for the user cold-start problem. We find that *Node2Vec* and *DeepWalk* provide the highest recommendation accuracy and user coverage [96] across all methods.

$\boxed{P7}$ presents our research on using variational autoencoders for session-based job recommendations. Specifically, to provide personalized job recommendations to users in a setting, in which we do not have full user preference histories available, we employ autoencoders to create latent representations of the limited preference information available in the anonymous user sessions (see Section 2.3.2). Our approach recommends jobs within new sessions by employing a *k*-nearest neighbor approach based on the inferred latent session representations generated via standard autoencoders [36], denoising autoencoders [274], and variational autoencoders [135]. Our evaluation results on session-based job recommendation datasets (e.g., based on XING from the RecSys challenge 2017 [12]) show that our approach based on variational autoencoders provides the most robust results compared to state-of-the-art methods such as *GRU4Rec* [113], *session-KNN* [124], or *sequential session-KNN* [182]. Here, we do not only evaluate recommendation accuracy, but also novelty metrics [272] such as system-based novelty (i.e., how unexplored is the recommended job in general [221]) and session-based novelty (i.e., how surprising is the recommended job for the current user session [289]). To further illustrate the usefulness of variational autoencoders for recommendations, in another paper [236], we utilize them to incorporate a user's country information into context-aware music recommendations. Specifically, we incorporate the users' country context into the variational autoencoder architecture via a gating mechanism. Our evaluation results show that our country- and context-aware recommendation approach provides higher recommendation accuracy than related baselines (e.g., variational autoencoders without country information [176]).

Contribution 4: Addressing Users' Privacy Constraints and the Trade-Off Between Accuracy and Privacy in Recommendations (2021-2023)

$\boxed{P8}$ studies the robustness of meta matrix factorization (*MetaMF*) against privacy constraints of users in recommender systems. For this, we conduct a reproducibility study of the original *MetaMF* paper [178], and investigate the sensitivity of this approach to the limited availability of user preference information, e.g., when users employ privacy constraints by hiding a certain part of their preferences from the system (see Section 2.3.1). Therefore, we deactivate the *meta learning* [271] component to evaluate the robustness of *MetaMF* against varying privacy constraints. Additionally, we study how users that differ in their profile size (i.e., number of ratings or implicit item preferences) are affected by varying privacy constraints. On the five datasets *Douban* [118], *Hetrec-MovieLens* [50], *MovieLens* 1M [108], *Ciao* [103], and *Jester* [99] (we share the dataset samples via *Zenodo* [208]), we demonstrate that *meta learning* is essential for *MetaMF*'s
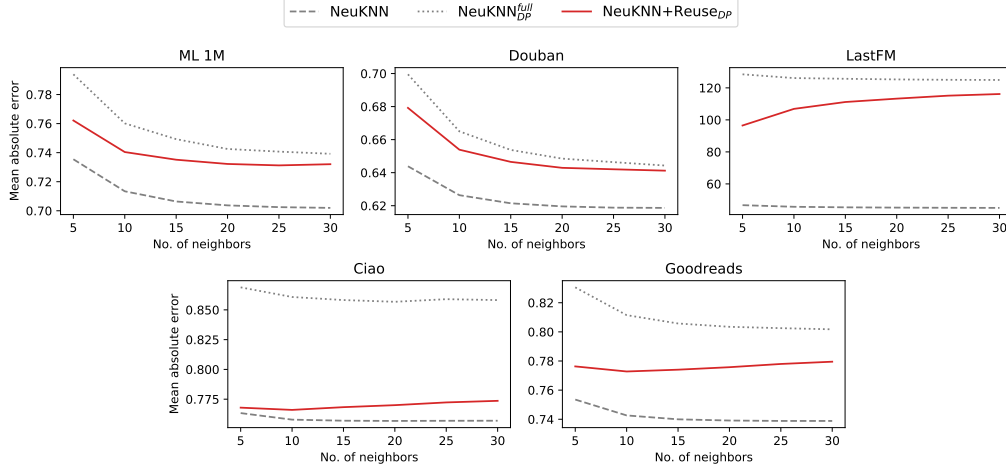
Figure 3.3: Mean absolute error (MAE) of neural-based *KNN* recommender system variants. Our results indicate that combining neighborhood reuse with *differential privacy* (*NeuKNN+Reuse$_{DP}$*) yields better accuracy (lower MAE) than neural-based methods that do not apply neighborhood reuse (*NeuKNN$_{DP}^{full}$*) [206].

robustness against users' privacy constraints. We also show that users with small profiles can afford a higher degree of privacy than users with large profiles.

[P9] addresses the accuracy-privacy trade-off in differentially-private recommender systems. Specifically, we propose our *ReuseKNN* recommendation approach, which aims to reduce the decrease in accuracy due to the application of *differential privacy* [74,76] on users' preference information [38]. We achieve this by identifying small but highly reusable neighborhoods for *k*-nearest neighbor-based recommendation approaches. Therefore, only this small set of users needs to be protected with *differential privacy*, and the majority of the users do not need to be protected, since they are rarely exploited as neighbors, i.e., they have a small privacy risk [180] as defined in Section 2.3.1. We find that with *ReuseKNN*, in the case of a Last.fm dataset sample, only 68.20% of the users need to be protected with *differential privacy*, while a traditional *UserKNN* approach [111] requires the protection of 99.89% of the users. We validate if this also leads to an improved accuracy-privacy trade-off in various recommendation settings. Figure 3.3 shows the recommendation accuracy results of neural-based CF approaches [109] when our neighborhood reuse concept is applied (*NeuKNN+Reuse$_{DP}$*, i.e., only vulnerable users are protected), and when it is not applied (*NeuKNN$_{DP}^{full}$*, i.e., all users are protected). Additionally, we include a baseline approach without any application of *differential privacy* (*NeuKNN*, i.e., no users are protected). We see that (i) *NeuKNN* provides the best accuracy results according to the mean absolute error [279], but without any privacy guarantees, (ii) *NeuKNN$_{DP}^{full}$* provides the worst accuracy results, but with the highest privacy guarantees, and (iii) that our *NeuKNN+Reuse$_{DP}$* approach provides a better accuracy-privacy trade-off than the other methods. Additionally, in this work, we outline connections between privacy, and item coverage [112], popularity bias [20,207], and fairness [80].

26

Finally, $\boxed{P10}$ further discusses the accuracy-privacy trade-off in differentially-private recommendations by surveying the literature in this field. Therefore, we identify 26 papers that apply *differential privacy* either (i) to the user representations (e.g., as we do it in $\boxed{P9}$), (ii) directly to the recommendation model updates (e.g., when calculating gradients locally), or (iii) after the recommendation model training process (e.g., applying noise to the trained user and item embeddings). We find that these papers address the accuracy-privacy trade-off in three different ways: (i) using auxiliary data to foster recommendation accuracy (e.g., incorporate preferences of other users), (ii) reducing the noise level that is needed (e.g., requiring the minimal amount of noise to still ensure *differential privacy*), and (iii) limit when to apply *differential privacy* (e.g., as we do it in $\boxed{P9}$).

3.3 Fairness and Popularity Bias in Recommender Systems

This section discusses our research on fairness and popularity bias in recommender systems. This contains four publications that study popularity bias for user groups that differ in mainstreaminess (i.e., users' inclination towards mainstream content [31]) and gender $\boxed{P11}$ $\boxed{P12}$ $\boxed{P13}$ $\boxed{P14}$ (*Contribution 5*). This section also describes two papers on understanding popularity bias mitigation and amplification using online and offline evaluation studies $\boxed{P15}$ $\boxed{P16}$ (*Contribution 6*). Another journal article analyzes the long-term dynamics of fairness (e.g., *individual* vs. *group fairness* trade-offs) in algorithmic decision support in a labor market setting using agent-based modeling techniques $\boxed{P17}$ (*Contribution 7*).

Contribution 5: Measuring Popularity Bias for User Groups Differing in Mainstreaminess and Gender (2020-2022)

$\boxed{P11}$ analyzes the unfairness of popularity bias in music recommendations. Specifically, we reproduce a study by Abdollahpouri et al. [7], in which the authors find that personalized recommendation algorithms in the movie domain are biased towards popular items, and that this popularity bias also leads to the unfair treatment of users with little interest into popular content (see Section 2.4.2). We conduct this reproducibility study in the music domain using a newly created dataset sample [142] gathered from Last.fm. Figure 3.4 shows that our results are in line with the ones of [7] since all evaluated recommendation algorithms tend to favor popular items also in the music domain. In the case of the *Most-Popular* algorithm, as expected, the strongest evidence for popularity bias can be found. In the case of traditional *UserKNN* [111] and *Non-negative Matrix Factorization* (*NMF*) [183], we also see a positive relationship between item (i.e., music artist) popularity and recommendation frequency. Finally, for *UserKNN* and *NMF*, we find that beyond-mainstream (*BeyMS*) users receive less accurate recommendations than mainstream (*MS*) users (see Figure 3.5a).

In $\boxed{P12}$, we analyze the unfairly treated *BeyMS* user group in more detail by identifying subgroups of beyond-mainstream music listeners. For this, we create a new dataset termed *LFM-BeyMS*, which contains (among others) audio features

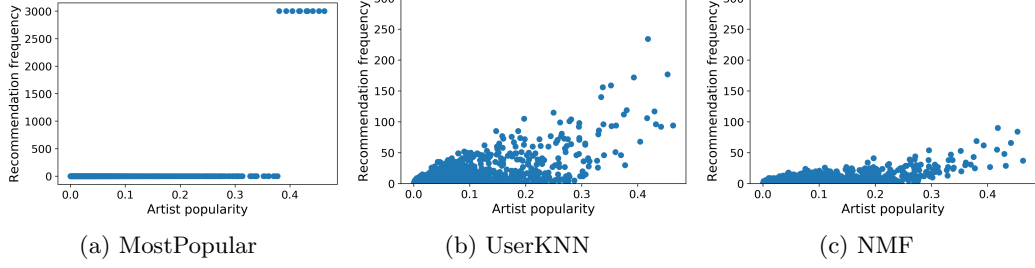| (a) MostPopular | (b) UserKNN | (c) NMF |

Figure 3.4: Correlation of music artist popularity and recommendation frequency. All three algorithms investigated tend to favor popular music artists [157].

of the music tracks listened to by more than 2,000 *BeyMS* users. Using these audio features and unsupervised clustering techniques, we identify four clusters of beyond-mainstream music and music listeners: (i) U_{folk}, listeners of music with high acousticness such as "folk", (ii) U_{hard}, listeners of high energy music such as "hardrock", (iii) U_{ambi}, listeners of music with high acousticness and instrumentalness such as "ambient", and (iv) U_{elec}, listeners of high energy music with high instrumentalness such as "electronica". Figure 3.5b shows that there is a substantial difference in recommendation accuracy between these subgroups of *BeyMS* users. While U_{ambi} users, on average, even receive better recommendation accuracy results than *MS* users, U_{hard} users receive the worst recommendation accuracy results. When relating our results to the openness of the subgroups' users towards music listened to by the other subgroups, we find that U_{ambi} is the most open group, while U_{hard} is the least open group. This is in line with related research [260], which has shown that a user's openness towards content consumed by other users is positively correlated with recommendation accuracy.

P13 studies if popularity bias in music recommender systems affect users of different genders in the same way. To answer this question, we analyze seven recommendation algorithms, *Random*, *MostPopular*, *ItemKNN* [234], *Sparse Linear Method* (*SLIM*) [211], *Alternating Least Squares Matrix Factorization* (*ALS*) [119], *Matrix Factorization with Bayesian Personalized Ranking* (*BPR*) [228], and *Variational Autoencoder for CF* (*VAE*) [176], on a Last.fm dataset sample based on the *LFM-2b* dataset [196, 237]. We find that all personalized recommendation algorithms investigated in this study, except for *SLIM*, intensify the popularity bias for female users. Thus, not only user groups differing in mainstreaminess, but also user groups differing in gender are affected differently by popularity bias.

Finally, P14 validates the findings of P11 and P12 in three additional multimedia domains, namely (i) movies (*MovieLens-1M* [108]), (ii) books (*BookCrossing* [290]), and (iii) animes (*MyAnimeList* [225]). For these datasets, we create dataset samples [144] with user groups that differ in their inclination to popular and mainstream content, and analyze popularity bias of various CF-based recommendation algorithms on the levels of items and users. On the item level, we find that the probability of an item to be recommended strongly correlates with the popularity of the item. On the user level, we find that users with the least inclination to popular content also receive the worst recommendation quality.

(a) MAE of *MS* and *BeyMS* users.
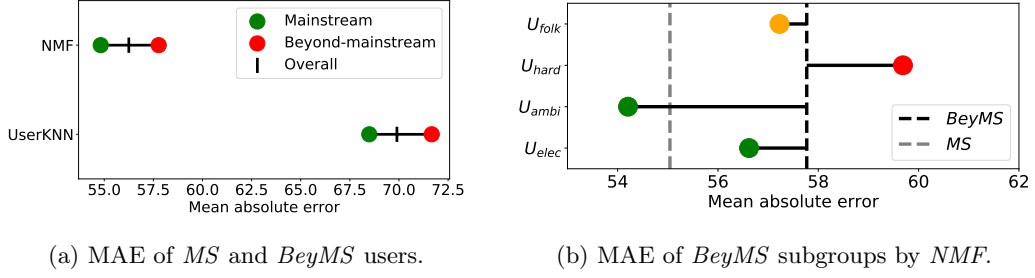

(b) MAE of *BeyMS* subgroups by *NMF*.

Figure 3.5: (a) Recommendation accuracy measured by the mean absolute error (MAE) of *NMF* and *UserKNN* for mainstream (*MS*) and beyond-mainstream (*BeyMS*) user groups in Last.fm: *BeyMS* users receive a substantially lower recommendation quality (i.e., higher MAE) compared to *MS* users. (b) Comparison of the MAE scores reached by *NMF* for the four *BeyMS* subgroups with the ones reached by *NMF* for *BeyMS* (black dashed line) and *MS* (grey dashed line). There are substantial differences between the subgroups in terms of MAE, especially when comparing U_{hard} with U_{ambi}, i.e., two subgroups differing in their openness to music listened to by users of other subgroups [154].

Contribution 6: Understanding Popularity Bias Mitigation and Amplification in Recommendations (2022-2023)

$\boxed{P15}$ presents an online study on popularity bias mitigation (see Section 2.4.2) in a news article recommendation setting. To conduct our online study, we collaborate with *DiePresse*, a popular Austrian online news platform, and discuss the introduction of personalized, content-based news article recommendations into the platform as a replacement for unpersonalized *MostPopular* recommendations. Our content-based recommendation algorithm [63, 181] is based on latent representations of news articles using *Latent Dirichlet Allocation (LDA)* [43]. We conducted our online study in a two-week time window (27th of October 2020 to 9th of November 2020), in which we tracked user preferences (i.e., clicks on news articles) of more than one Million anonymous user sessions, and more than 15,000 signed in (subscribed) users of *DiePresse*. Within our two-week online study, also two significant events happened that could influence the reading behavior of users: (i) the COVID-19 lockdown announcements in Austria on the 31st of October 2020, and (ii) the Vienna terror attack on the 2nd of November 2020.

Figure 3.6 shows the results of our online study in terms of *skewness* and *kurtosis* of the news article popularity distribution (i.e., number of article reads) across the two weeks, and for both user groups (i.e., anonymous and subscribed users). Here, *skewness* measures the asymmetry, and *kurtosis* measures the "tailedness" of the popularity distribution [34]. For both metrics, high values indicate a popularity biased news consumption, which could lead to filter bubble and echo chamber effects [89]. At the beginning of the online study, where *MostPopular* recommendations were shown, we see a large gap between the two user groups: while anonymous users mainly read popular news articles, and thus, are prone to popularity bias, subscribed users show a much more balanced reading behavior. At
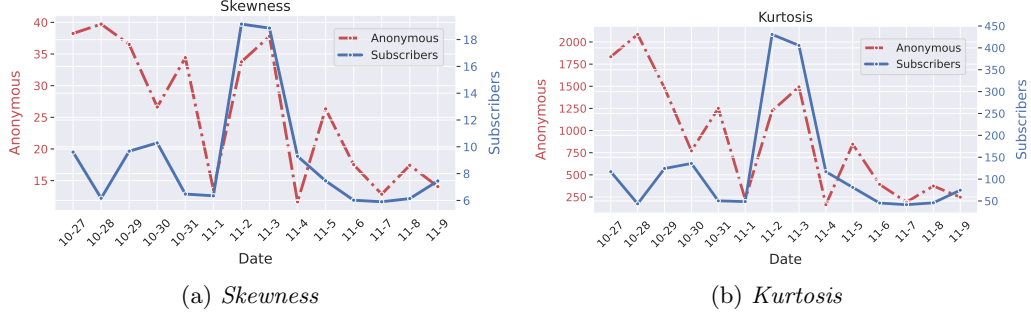
29

(a) *Skewness*　　　　　　　　　　　(b) *Kurtosis*

Figure 3.6: Mitigation of popularity bias in news article consumption, measured by (a) *skewness* and (b) *kurtosis* based on the number of article reads for each day of our two-week online study. At the beginning of the study, the *MostPopular* news article recommendations were replaced by personalized, content-based recommendations. We find that popularity bias can be mitigated by introducing personalized news article recommendations in the case of anonymous users [163].

the end of the study, i.e., after two weeks of personalized recommendations, we see a considerably smaller difference between the two user groups, which means that the introduction of personalized, content-based news article recommendations helped to mitigate popularity bias in the case of anonymous users already after two weeks. However, in the case of significant events, e.g., the Vienna terror attack on the 2nd of November 2020, both user groups are mostly interested into popular articles reporting on the particular event. In another work [203], we also find that content-based recommendations can help to mitigate popularity bias in the case of recommendations provided in a data and algorithm sharing platform.

In $\boxed{P16}$, we analyze miscalibration [177,255] and popularity bias amplification (in terms of the popularity lift metric [7,8]) in music, movie, and anime recommender systems. For this, we extend the *MovieLens 1M* [108], *LFM-1b* [235,238], and *MyAnimeList* [225] datasets with genre information of the items, and publish these new dataset samples via *Zenodo* [143]. Then we measure accuracy, miscalibration, and popularity bias amplification (i.e., popularity lift) for various recommendation algorithms (e.g., *NMF* [183] and *co-clustering*-based CF [98]), and for user groups differing in their inclination to popular and mainstream content, i.e., (i) *LowPop* (low interest in popular content), (ii) *MedPop* (medium interest in popular content), and (iii) *HighPop* (high interest in popular content). We find that there is a connection between these three metrics, since the *LowPop* user group, which receives the worst recommendation accuracy results, is also the user group, which receives the most miscalibrated and popularity biased recommendations. Finally, we investigate to what extent particular genres contribute to the inconsistency of recommendation performance in terms of miscalibration and popularity bias amplification. We find that there are indeed genres that highly contribute to inconsistent and popularity biased recommendation results. One example is the "Hentai" genre in the case of our *MyAnimeList* dataset sample: this is a genre, which is highly popular for a specific user group (i.e., *LowPop*), and unpopular for the other user groups (i.e., *MedPop* and *HighPop*).

Contribution 7: Studying Long-Term Dynamics of Fairness in Algorithmic Decision Support (2022-2023)

$\boxed{P17}$ studies the long-term dynamics of fairness in algorithmic decision support (see Section 2.4.1) in a labor market setting [94]. Specifically, we develop and evaluate an agent-based simulation model to investigate the impact of decisions caused by a public employment service that decides which jobseekers receive targeted help using a decision support tool. This tool uses a logistic regression model [280] to classify jobseekers into low- and high-prospects. We use synthetic data that describes a pool of jobseekers with unevenly distributed skills between two groups that differ with respect to a protected attribute. We test two variants of our prediction model: (i) a biased version that augments knowledge about the actual skills of a jobseeker with knowledge about the protected attribute, and (ii) an unbiased version that solely relies on the skills of a jobseeker. Based on the classification into low-prospects and high-prospects, our agent-based simulation model updates the skills of the jobseekers after each iteration accordingly (e.g., a high-prospect receives help, and thus also the skills of this jobseeker increase).

Our results show that there is a trade-off between different long-term fairness goals. On the one hand, when using the biased prediction model, the inequality between the two protected groups is reduced at the end of the simulation. This means, that *statistical parity* in the dataset [30] increases, and that the system is fair from a *group fairness* perspective. However, on the other hand, the number of misclassifications of jobseekers in the unprivileged group increases: some jobseekers are classified as low-prospect mainly because of their sensitive attribute, although they should belong to the high-prospect group. This means that the system is unfair from an *individual fairness* perspective. Although this study was not conducted in the field of recommender systems, we believe that the applied method (i.e., agent-based modeling) could also be of use when studying long-term fairness dynamics of recommender systems. Additionally, our findings with respect to the trade-off between *individual* and *group fairness* are also highly relevant for the research area of fair recommender systems.

3.4 Summary of Contributions and Reproducibility of Research Results

This section summarizes the 7 scientific contributions described in the previous sections. Additionally, the reproducibility of the findings are discussed.

List of Contributions

1. **Using cognitive models for a transparent design and implementation process of recommender systems (2018-2021)**: we propose a tag recommendation approach based on a model of human episodic memory $\boxed{P1}$, and two music recommendation approaches based on activation process in human memory $\boxed{P2}$ $\boxed{P3}$. Additionally, we identify three types of psychology-informed recommender systems: (i) cognition-inspired, (ii) personality-aware, and (iii) affect-aware recommender systems $\boxed{P4}$.

2. **Illustrating to what extent components of the cognitive model ACT-R contribute to recommendations (2022-2023)**: we illustrate to what extent components of ACT-R (e.g., *BLL* or *valuation*) have contributed to the generation of music recommendation lists. Based on this, explanations for the music recommendation could be derived $\boxed{P5}$.

3. **Addressing limited user preference information in cold-start and session-based recommendation settings (2018-2020)**: we model a user's trust network using regular equivalence to address the user cold-start problem $\boxed{P6}$. Additionally, we demonstrate the usefulness of variational autoencoders for session-based job recommendations $\boxed{P7}$.

4. **Addressing users' privacy constraints and the accuracy privacy trade-off in recommendations (2021-2023)**: we study privacy constraints of users (e.g., hiding preferences) in meta matrix factorization $\boxed{P8}$, design a neighborhood reuse approach $\boxed{P9}$, and survey the literature for differentially-private collaborative filtering recommender systems $\boxed{P10}$.

5. **Measuring popularity bias for user groups differing in mainstreaminess and gender (2020-2022)**: we study popularity bias $\boxed{P11}$, characteristics of beyond-mainstream users $\boxed{P12}$, and differences with respect to users' gender in music recommendations $\boxed{P13}$. We also show the presence of popularity bias in movie, book, and anime recommendations $\boxed{P14}$.

6. **Understanding popularity bias mitigation and amplification in recommendations (2022-2023)**: we analyze and mitigate popularity in news article recommender systems $\boxed{P15}$, and study to what extent recommendations amplify popularity bias in the music, movie, and anime domains $\boxed{P16}$.

7. **Studying long-term dynamics of fairness in algorithmic decision support (2022-2023)**: we show the usefulness of agent-based modeling techniques for studying long-term dynamics of algorithmic fairness in a labor market setting. Additionally, we find evidence for the presence of the trade-off between *individual* and *group fairness* in this setting $\boxed{P17}$.

Reproducibility of Research Results

To foster the reproducibility of these research results and findings, we provide information on the used source-code and dataset samples in all publications. In cases, in which we create new dataset samples or implement novel recommendation pipelines, we make them freely available via *Zenodo* or *GitHub*. For example, to implement and evaluate our cognitive-inspired recommendation approaches, we build upon our *TagRec* framework [146,148,266], and extend it with music recommendation approaches. Another example is our *Last.fm user group* dataset sample [142] that can be used to study fairness and popularity bias in recommender systems. Additionally, we contribute to reproducibility studies by presenting two papers enlisted in this habilitation in the reproducibility track of the *European Conference on Information Retrieval (ECIR'2020 and ECIR'2021)* $\boxed{P8}$ $\boxed{P11}$.

A list of the new dataset samples and recommendation pipelines created in the publications that are part of this habilitation is given in the following:

1. The *TagRec* framework [146, 148, 266] used to design, develop, and evaluate cognitive-inspired algorithms for tag and music recommendations: `https://github.com/learning-layers/TagRec`.

2. A *GitHub* repository with the material to generate sequential music recommendations and to illustrate to what extent the components of the cognitive model ACT-R contribute to the generation of music recommendation lists [201]: `https://github.com/hcai-mms/actr`.

3. A dataset sample based on the *LFM-2b* dataset [196, 237] used to generate and evaluate sequential music recommendations [200]. This *Zenodo* repository also contains the pre-calculated embeddings for the *BPR* approach.

4. Source-code and dataset references for using variational autoencoders in the setting of session-based job recommendations [167]: `https://github.com/lacic/session-knn-ae`. This *GitHub* repository also contains implementations of beyond-accuracy evaluation metrics (e.g., diversity and novelty) for session-based recommender systems.

5. A dataset for studying privacy constraints of different users groups using meta matrix factorization [208] accompanied by a *GitHub* repository: `https://github.com/pmuellner/RobustnessOfMetaMF`.

6. The material for the differentially-private *ReuseKNN* [206] recommender system: `https://github.com/pmuellner/ReuseKNN`. This *GitHub* repository also contains the implementation of *Neural CF*, as well as source-code for sampling user preference histories in the datasets.

7. A dataset for studying beyond-mainstream users in music recommender systems [209] accompanied by a *GitHub* repository: `https://github.com/pmuellner/supporttheunderground`. Apart from popularity bias evaluation metrics, this *GitHub* repository contains implementations of unsupervised clustering techniques to analyze audio features of music tracks.

8. Datasets containing different user groups to study fairness and popularity bias in music, movie, book, and anime recommender systems [142, 144]. For calculating calibration-based metrics in these settings, an extended version of these datasets also contains genre information for the items [143].

9. A Python-based pipeline to process the datasets used in [147, 153, 157] for studying fairness and popularity bias in recommender systems: `https://github.com/domkowald/FairRecSys`. This *GitHub* repository can also be used as a basis to develop popularity bias mitigation methods.

By publishing these resources, the author of this habilitation hopes to contribute to reproducible research practices in the field of recommender systems. As mentioned already in Section 2.1, the reproducibility of research results is highly important for being able to track progress in recommender systems research.

Chapter 4

Outlook and Future Research

This chapter gives an outlook into future research directions of this habilitation.

Transparency and Cognitive Models in Recommender Systems

The underlying algorithms of modern recommender systems are often based on purely data-driven machine learning models. Although these approaches provide high accuracy, they are based on principles of artificial intelligence rather than human intelligence. One consequence could be that the logic of these models is not directly understandable by humans, which could lead to non-transparent algorithmic decisions [250]. This habilitation has shown that using psychological theories, and modeling the underlying cognitive processes that describe how humans access information in their memory, is one way to overcome this issue, and at the same time, to generate accurate recommendations (see Section 3.1).

Besides MINERVA2 [114], the cognitive architecture ACT-R [22] provides an excellent basis by formalizing two kinds of human memory: (i) declarative memory, and (ii) procedural memory. The declarative memory corresponds to things that humans know by determining the importance of information chunks, while the procedural memory corresponds to knowledge of how humans do things by defining production rules for making decisions. This habilitation has focused on modeling declarative memory processes for a transparent design process of cognitive-inspired recommender systems. Thus, in future research, I aim to investigate to what extent also the procedural memory module of ACT-R can be used to design recommendation models (e.g., by adapting the *SNIF-ACT* [93] user navigation model). Here, one interesting research question would be if the defined production rules could further contribute to transparency aspects of cognitive-inspired recommender systems. This question could be answered by conducting user studies following well-established procedures in the field (e.g., [139,216,263]).

Privacy and Limited Preference Information in Recommender Systems

Privacy is a key requirement for recommender systems, since there are multiple privacy threats to users in these systems. For example, disclosing users' preference information to untrusted third parties [49], or inferring users' sensitive attributes such as gender [287]. Privacy is also related to the issue of limited availability

of user preference information, since users increasingly care about their privacy and may not want to share their preferences with the system [138, 194, 270]. Additionally, initiatives such as the *European General Data Protection Regulation (GDPR)* restrict the use of user preference information to generate recommendations [40, 62]. This habilitation has addressed session-based and cold-start recommendation settings, and the accuracy-privacy trade-off when applying *differential privacy* to the users' preference information (see Section 3.2).

In the future, I plan to not only study the trade-off between accuracy and privacy, but also to investigate other relevant trade-offs between recommendation objectives. This includes the trade-off between privacy and fairness [80]. Here, an interesting research question would be if different user groups are treated differently by the accuracy drops due to privacy-preserving technologies, such as *differential privacy*. For this, related studies from the field of private and fair machine learning (e.g., [29]) could be adapted for recommender systems. Additionally, studying privacy dynamics in recommendations using agent-based simulations would be a promising research direction, as described in our position paper [204] presented in the *SimuRec* workshop of *ACM RecSys* 2021.

Fairness and Popularity Bias in Recommender Systems

Biases in the perception and behavior of humans are captured, reflected, and potentially amplified in recommender systems [55, 92, 168]. The replication of popularity bias is a common issue in collaborative filtering-based recommender systems, which leads to the overrepresentation of popular items in the recommendation lists. The research presented in this habilitation has shown that users with little interest in popular content receive worse recommendation accuracy than users that like to consume popular content. Based on this, these users are treated in an unfair way by the recommender system (see Section 3.3).

In my future research, I plan to work on popularity bias mitigation methods to reduce the accuracy differences between the user groups, and with this, increase the fairness in the system. For this, not only technical debiasing methods (e.g., in- or post-processing [9]), but also novel multidisciplinary approaches using models from psychology and physics should be developed. For the former, ACT-R [22] could be a promising basis to build strongly personalized user models, and for the latter, techniques from physics-informed machine learning could be transferred to fairness problems, as described in our recent *arXiv* pre-print [242].

Reproducibility Aspects of this Habilitation

I want to highlight the importance of reproducibility for the research field of recommender systems [32, 87]. This habilitation has provided several resources to foster the reproducibility of the presented research results (see Section 3.4). In the future, I want to further contribute to the reproducibility of machine learning research in general, and recommender systems research in particular, by discussing barriers and best practices as outlined in our recent publications [162, 247].

Finally, I hope that the scientific results and findings of this habilitation contribute to advancing research on the trustworthiness of recommender systems.

Bibliography

[1] Himan Abdollahpouri, Gediminas Adomavicius, Robin Burke, Ido Guy, Dietmar Jannach, Toshihiro Kamishima, Jan Krasnodebski, and Luiz Pizzato. Multistakeholder recommendation: Survey and research directions. *User Modeling and User-Adapted Interaction*, 30:127–158, 2020.

[2] Himan Abdollahpouri, Robin Burke, and Bamshad Mobasher. Controlling popularity bias in learning-to-rank recommendation. In *Proceedings of the Eleventh ACM Conference on Recommender Systems*, pages 42–46, 2017.

[3] Himan Abdollahpouri, Robin Burke, and Bamshad Mobasher. Recommender systems as multistakeholder environments. In *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*, pages 347–348, 2017.

[4] Himan Abdollahpouri, Robin Burke, and Bamshad Mobasher. Managing popularity bias in recommender systems with personalized re-ranking. In *The Thirty-Second International FLAIRS Conference*, 2019.

[5] Himan Abdollahpouri, Edward C Malthouse, Joseph A Konstan, Bamshad Mobasher, and Jeremy Gilbert. Toward the next generation of news recommender systems. In *Companion Proceedings of the Web Conference 2021*, pages 402–406, 2021.

[6] Himan Abdollahpouri, Masoud Mansoury, Robin Burke, and Bamshad Mobasher. The impact of popularity bias on fairness and calibration in recommendation. *arXiv preprint arXiv:1910.05755*, 2019.

[7] Himan Abdollahpouri, Masoud Mansoury, Robin Burke, and Bamshad Mobasher. The unfairness of popularity bias in recommendation. In *RMSE Workshop co-located with the 13th ACM Conference on Recommender Systems (RecSys 2019)*, 2019.

[8] Himan Abdollahpouri, Masoud Mansoury, Robin Burke, and Bamshad Mobasher. The connection between popularity bias, calibration, and fairness in recommendation. In *Fourteenth ACM Conference on Recommender Systems*, pages 726–731, 2020.

[9] Himan Abdollahpouri, Masoud Mansoury, Robin Burke, Bamshad Mobasher, and Edward Malthouse. User-centered evaluation of popularity

bias in recommender systems. In *Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*, pages 119–129, 2021.

[10] Fabian Abel. We know where you should work next summer: Job recommendations. In *Proceedings of the 9th ACM Conference on Recommender Systems*, pages 230–230, 2015.

[11] Fabian Abel, András Benczúr, Daniel Kohlsdorf, Martha Larson, and Róbert Pálovics. Recsys challenge 2016: Job recommendations. In *Proceedings of the 10th ACM Conference on Recommender Systems*, pages 425–426, 2016.

[12] Fabian Abel, Yashar Deldjoo, Mehdi Elahi, and Daniel Kohlsdorf. Recsys challenge 2017: Offline and online evaluation. In *Proceedings of the Eleventh ACM Conference on Recommender Systems*, pages 372–373, 2017.

[13] Panagiotis Adamopoulos and Alexander Tuzhilin. On unexpectedness in recommender systems: Or how to better expect the unexpected. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(4):1–32, 2014.

[14] Gediminas Adomavicius, Konstantin Bauman, Alexander Tuzhilin, and Moshe Unger. Context-aware recommender systems: From foundations to recent developments context-aware recommender systems. In *Recommender Systems Handbook*, pages 211–250. Springer, 2021.

[15] Gediminas Adomavicius, Jesse C Bockstedt, Shawn P Curley, and Jingjing Zhang. Do recommender systems manipulate consumer preferences? a study of anchoring effects. *Information Systems Research*, 24(4):956–975, 2013.

[16] Gediminas Adomavicius and YoungOk Kwon. Improving aggregate recommendation diversity using ranking-based techniques. *IEEE Transactions on Knowledge and Data Engineering*, 24(5):896–911, 2011.

[17] Gediminas Adomavicius, Ramesh Sankaranarayanan, Shahana Sen, and Alexander Tuzhilin. Incorporating contextual information in recommender systems using a multidimensional approach. *ACM Transactions on Information Systems (TOIS)*, 23(1):103–145, 2005.

[18] Gediminas Adomavicius and Alexander Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6):734–749, 2005.

[19] Gediminas Adomavicius and Alexander Tuzhilin. Context-aware recommender systems. In *Recommender Systems Handbook*, pages 217–253. Springer, 2010.

[20] Abdul Basit Ahanger, Syed Wajid Aalam, Muzafar Rasool Bhat, and Assif Assad. Popularity bias in recommender systems - a review. In *International Conference on Emerging Technologies in Computer Engineering*, pages 431–444. Springer, 2022.

[21] Amr Ahmed, Nino Shervashidze, Shravan Narayanamurthy, Vanja Josifovski, and Alexander J Smola. Distributed large-scale natural graph factorization. In *Proceedings of the 22nd International Conference on World Wide Web*, pages 37–48, 2013.

[22] John R Anderson, Daniel Bothell, Michael D Byrne, Scott Douglass, Christian Lebiere, and Yulin Qin. An integrated theory of the mind. *Psychological Review*, 111(4):25 pages, 2004.

[23] John R Anderson, Michael Matessa, and Christian Lebiere. Act-r: A theory of higher level cognition and its relation to visual attention. *Human–Computer Interaction*, 12(4):439–462, 1997.

[24] John R Anderson and Lael J Schooler. Reflections of the environment in memory. *Psychological Science*, 2(6):396–408, 1991.

[25] Vito Walter Anelli, Yashar Deldjoo, Tommaso Di Noia, Antonio Ferrara, and Fedelucio Narducci. Federank: User controlled feedback with federated recommender systems. In *Advances in Information Retrieval: 43rd European Conference on IR Research (ECIR 2021)*, pages 32–47. Springer, 2021.

[26] Arda Antikacioglu and R Ravi. Post processing recommender systems for diversity. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 707–716, 2017.

[27] Müslüm Atas, Alexander Felfernig, Seda Polat-Erdeniz, Andrei Popescu, Thi Ngoc Trang Tran, and Mathias Uta. Towards psychology-aware preference construction in recommender systems: Overview and research issues. *Journal of Intelligent Information Systems*, 57:467–489, 2021.

[28] Ricardo Baeza-Yates and Berthier Ribeiro-Neto. *Modern information retrieval*, volume 463. ACM Press New York, 1999.

[29] Eugene Bagdasaryan, Omid Poursaeed, and Vitaly Shmatikov. Differential privacy has disparate impact on model accuracy. *Advances in Neural Information Processing Systems*, 32, 2019.

[30] Solon Barocas and Andrew D Selbst. Big data's disparate impact. *California Law Review*, pages 671–732, 2016.

[31] Christine Bauer and Markus Schedl. Global and country-specific mainstreaminess measures: Definitions, analysis, and usage for improving personalized music recommendation systems. *PloS One*, 14(6):e0217389, 2019.

[32] Joeran Beel, Corinna Breitinger, Stefan Langer, Andreas Lommatzsch, and Bela Gipp. Towards reproducibility in recommender-systems research. *User Modeling and User-Adapted Interaction*, 26:69–101, 2016.

[33] Ghazaleh Beigi and Huan Liu. A survey on privacy in social media: Identification, mitigation, and applications. *ACM Transactions on Data Science*, 1(1):1–38, 2020.

[34] Alejandro Bellogín, Pablo Castells, and Iván Cantador. Statistical biases in information retrieval metrics for recommender systems. *Information Retrieval Journal*, 20:606–634, 2017.

[35] Jacob Benesty, Jingdong Chen, Yiteng Huang, and Israel Cohen. Pearson correlation coefficient. In *Noise Reduction in Speech Processing*, pages 37–40. Springer, 2009.

[36] Yoshua Bengio, Pascal Lamblin, Dan Popovici, and Hugo Larochelle. Greedy layer-wise training of deep networks. *Advances in Neural Information Processing Systems*, 19, 2006.

[37] Richard Berk, Hoda Heidari, Shahin Jabbari, Michael Kearns, and Aaron Roth. Fairness in criminal justice risk assessments: The state of the art. *Sociological Methods & Research*, 1:42, 2018.

[38] Shlomo Berkovsky, Tsvi Kuflik, and Francesco Ricci. The impact of data obfuscation on the accuracy of collaborative filtering. *Expert Systems with Applications*, 39(5):5033–5042, 2012.

[39] James R Bettman and C Whan Park. Effects of prior knowledge and experience and phase of the choice process on consumer decision processes: A protocol analysis. *Journal of Consumer Research*, 7(3):234–248, 1980.

[40] Asia J Biega, Peter Potash, Hal Daumé, Fernando Diaz, and Michèle Finck. Operationalizing the legal principle of data minimization for personalization. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 399–408, 2020.

[41] Vincent Bindschaedler, Reza Shokri, and Carl A Gunter. Plausible deniability for privacy-preserving data synthesis. *Proceedings of the VLDB Endowment*, 10(5), 2017.

[42] Reuben Binns. On the apparent conflict between individual and group fairness. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pages 514–524, 2020.

[43] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3(Jan):993–1022, 2003.

[44] Erik Brynjolfsson, Yu Jeffrey Hu, and Michael D Smith. From niches to riches: Anatomy of the long tail. *Sloan Management Review*, 47(4):67–71, 2006.

[45] Robin Burke. Hybrid recommender systems: Survey and experiments. *User Modeling and User-adapted Interaction*, 12:331–370, 2002.

[46] Robin Burke. Hybrid web recommender systems. *The Adaptive Web: Methods and Strategies of Web Personalization*, pages 377–408, 2007.

[47] Robin Burke, Alexander Felfernig, and Mehmet H Göker. Recommender systems: An overview. *AI Magazine*, 32(3):13–18, 2011.

[48] Robin Burke, Nasim Sonboli, and Aldo Ordonez-Gauger. Balanced neighborhoods for multi-sided fairness in recommendation. In *Conference on Fairness, Accountability and Transparency*, pages 202–214. PMLR, 2018.

[49] Joseph A Calandrino, Ann Kilzer, Arvind Narayanan, Edward W Felten, and Vitaly Shmatikov. You might also like: Privacy risks of collaborative filtering. In *Proceedings of 2011 IEEE Symposium on Security and Privacy*, pages 231–246. IEEE, 2011.

[50] Iván Cantador, Peter Brusilovsky, and Tsvi Kuflik. Second workshop on information heterogeneity and fusion in recommender systems (hetrec2011). In *Proceedings of the Fifth ACM Conference on Recommender Systems*, pages 387–388, 2011.

[51] Pablo Castells, Neil Hurley, and Saul Vargas. Novelty and diversity in recommender systems. In *Recommender Systems Handbook*, pages 603–646. Springer, 2021.

[52] Pablo Castells and Dietmar Jannach. Recommender systems: A primer. *arXiv preprint arXiv:2302.02579*, 2023.

[53] Pablo Castells and Alistair Moffat. Offline recommender system evaluation: Challenges and new directions. *AI Magazine*, 43(2):225–238, 2022.

[54] Òscar Celma and Perfecto Herrera. A new approach to evaluating novel recommendations. In *Proceedings of the 2008 ACM Conference on Recommender Systems*, pages 179–186, 2008.

[55] Jiawei Chen, Hande Dong, Xiang Wang, Fuli Feng, Meng Wang, and Xiangnan He. Bias and debias in recommender system: A survey and future directions. *ACM Transactions on Information Systems*, 41(3):1–39, 2023.

[56] Li Chen, Yonghua Yang, Ningxia Wang, Keping Yang, and Quan Yuan. How serendipity improves user satisfaction with recommendations? a large-scale user evaluation. In *The World Wide Web Conference*, pages 240–250, 2019.

[57] Xiaocong Chen, Lina Yao, Julian McAuley, Guanglin Zhou, and Xianzhi Wang. Deep reinforcement learning in recommender systems: A survey and new perspectives. *Knowledge-Based Systems*, 264:110335, 2023.

[58] Pern Hui Chia and Georgios Pitsilis. Exploring the use of explicit trust links for filtering recommenders: A study on epinions.com. *Journal of Information Processing*, 19:332–344, 2011.

[59] European Commission. Proposal for a regulation laying down harmonised rules on artificial intelligence (eu ai act), url: https://eur-lex.europa.eu/legal-content/en/txt/?uri=celex:52021pc0206, 2021. Accessed on November 22nd, 2023.

[60] Paolo Cremonesi and Dietmar Jannach. Progress in recommender systems research: Crisis? what crisis? *AI Magazine*, 42(3):43–54, 2021.

[61] Paolo Cremonesi, Yehuda Koren, and Roberto Turrin. Performance of recommender algorithms on top-n recommendation tasks. In *Proceedings of the Fourth ACM Conference on Recommender Systems*, pages 39–46, 2010.

[62] Rachel Cummings and Deven Desai. The role of differential privacy in gdpr compliance. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT'18)*, page 20, 2018.

[63] Marco De Gemmis, Pasquale Lops, Cataldo Musto, Fedelucio Narducci, and Giovanni Semeraro. Semantics-aware content-based recommender systems. *Recommender Systems Handbook*, pages 119–159, 2015.

[64] Gabriel de Souza Pereira Moreira, Felipe Ferreira, and Adilson Marques da Cunha. News session-based recommendations using deep neural networks. In *Proceedings of the 3rd Workshop on Deep Learning for Recommender Systems*, pages 15–23, 2018.

[65] Yashar Deldjoo, Dietmar Jannach, Alejandro Bellogin, Alessandro Difonzo, and Dario Zanzonelli. Fairness in recommender systems: Research landscape and future directions. *User Modeling and User-Adapted Interaction*, pages 1–50, 2023.

[66] Yashar Deldjoo, Markus Schedl, Paolo Cremonesi, and Gabriella Pasi. Recommender systems leveraging multimedia content. *ACM Computing Surveys (CSUR)*, 53(5):1–38, 2020.

[67] Sebastian Dennerlein, Dieter Theiler, Peter Marton, Patricia Santos Rodriguez, John Cook, Stefanie Lindstaedt, and Elisabeth Lex. Knowbrain: An online social knowledge repository for informal workplace learning. In *10th European Conference on Technology Enhanced Learning (EC-TEL 2015)*, pages 509–512. Springer, 2015.

[68] Christian Desrosiers and George Karypis. A comprehensive survey of neighborhood-based recommendation methods. *Recommender Systems Handbook*, pages 107–144, 2010.

[69] Tommaso Di Noia, Nava Tintarev, Panagiota Fatourou, and Markus Schedl. Recommender systems under european ai regulations. *Communications of the ACM*, 65(4):69–73, 2022.

[70] Gideon Dror, Noam Koenigstein, Yehuda Koren, and Markus Weimer. The yahoo! music dataset and kdd-cup'11. In *Proceedings of KDD Cup 2011*, pages 3–18. PMLR, 2012.

[71] Tomislav Duricic, Hussain Hussain, Emanuel Lacic, Dominik Kowald, Denis Helic, and Elisabeth Lex. Empirical comparison of graph embeddings for trust-based collaborative filtering. In *Foundations of Intelligent Systems: 25th International Symposium, ISMIS 2020*, pages 181–191. Springer, 2020.

[72] Tomislav Duricic, Dominik Kowald, Emanuel Lacic, and Elisabeth Lex. Beyond-accuracy: A review on diversity, serendipity and fairness in recommender systems based on graph neural networks. *Frontiers in Big Data*, 6, 2023.

[73] Tomislav Duricic, Emanuel Lacic, Dominik Kowald, and Elisabeth Lex. Trust-based collaborative filtering: Tackling the cold start problem using regular equivalence. In *Proceedings of the 12th ACM Conference on Recommender Systems*, pages 446–450, 2018.

[74] Cynthia Dwork. Differential privacy: A survey of results. In *International Conference on Theory and Applications of Models of Computation*, pages 1–19. Springer, 2008.

[75] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. Fairness through awareness. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, pages 214–226, 2012.

[76] Cynthia Dwork, Aaron Roth, et al. The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3-4):211–407, 2014.

[77] Magdalini Eirinaki, Jerry Gao, Iraklis Varlamis, and Konstantinos Tserpes. Recommender systems for large-scale social networks: A review of challenges and solutions. *Future Generation Computer Systems*, 78:413–418, 2018.

[78] Magdalini Eirinaki, Malamati D Louta, and Iraklis Varlamis. A trust-aware system for personalized user recommendations in social networks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 44(4):409–421, 2013.

[79] Michael D Ekstrand, Anubrata Das, Robin Burke, Fernando Diaz, et al. Fairness in information access systems. *Foundations and Trends in Information Retrieval*, 16(1-2):1–177, 2022.

[80] Michael D Ekstrand, Rezvan Joshaghani, and Hoda Mehrpouyan. Privacy for all: Ensuring fair and equitable privacy protections. In *Conference on Fairness, Accountability and Transparency*, pages 35–47. PMLR, 2018.

[81] Michael D Ekstrand, John T Riedl, Joseph A Konstan, et al. Collaborative filtering recommender systems. *Foundations and Trends in Human–Computer Interaction*, 4(2):81–173, 2011.

[82] Michael D Ekstrand, Mucun Tian, Ion Madrazo Azpiazu, Jennifer D Ekstrand, Oghenemaro Anuyah, David McNeill, and Maria Soledad Pera. All the cool kids, how do they fit in?: Popularity and demographic biases in recommender evaluation and effectiveness. In *Conference on Fairness, Accountability and Transparency*, pages 172–186. PMLR, 2018.

[83] Mehdi Elahi, Danial Khosh Kholgh, Mohammad Sina Kiarostami, Sorush Saghari, Shiva Parsa Rad, and Marko Tkalčič. Investigating the impact of recommender systems on user-based and item-based popularity bias. *Information Processing & Management*, 58(5):102655, 2021.

[84] Wenqi Fan, Xiangyu Zhao, Xiao Chen, Jingran Su, Jingtong Gao, Lin Wang, Qidong Liu, Yiqi Wang, Han Xu, Lei Chen, et al. A comprehensive survey on trustworthy recommender systems. *arXiv preprint arXiv:2209.10117*, 2022.

[85] Wenqi Fan, Xiangyu Zhao, Lin Wang, Xiao Chen, Jingtong Gao, Qidong Liu, and Shijie Wang. Trustworthy recommender systems: Foundations and frontiers. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 5796–5797, 2023.

[86] Alexander Felfernig, Gerhard Friedrich, Bartosz Gula, Martin Hitz, Thomas Kruggel, Gerhard Leitner, Rudolf Melcher, Daniela Riepan, Sabine Strauss, Erich Teppan, et al. Persuasive recommendation: Serial position effects in knowledge-based recommender systems. In *Persuasive Technology: Second International Conference on Persuasive Technology (PERSUASIVE 2007)*, pages 283–294. Springer, 2007.

[87] Maurizio Ferrari Dacrema, Simone Boglio, Paolo Cremonesi, and Dietmar Jannach. A troubling analysis of reproducibility and progress in recommender systems research. *ACM Transactions on Information Systems (TOIS)*, 39(2):1–49, 2021.

[88] Maurizio Ferrari Dacrema, Paolo Cremonesi, and Dietmar Jannach. Are we really making much progress? a worrying analysis of recent neural recommendation approaches. In *Proceedings of the 13th ACM conference on recommender systems*, pages 101–109, 2019.

[89] Seth Flaxman, Sharad Goel, and Justin M Rao. Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, 80(S1):298–320, 2016.

[90] Arik Friedman, Shlomo Berkovsky, and Mohamed Ali Kaafar. A differential privacy framework for matrix factorization recommender systems. *User Modeling and User-Adapted Interaction*, 26:425–458, 2016.

[91] Arik Friedman, Bart P Knijnenburg, Kris Vanhecke, Luc Martens, and Shlomo Berkovsky. Privacy aspects of recommender systems. *Recommender Systems Handbook*, pages 649–688, 2015.

[92] Batya Friedman and Helen Nissenbaum. Bias in computer systems. *ACM Transactions on Information Systems (TOIS)*, 14(3):330–347, 1996.

[93] Wai-Tat Fu and Peter Pirolli. Snif-act: A cognitive model of user navigation on the world wide web. *Human–Computer Interaction*, 22(4):355–412, 2007.

[94] Simon Gächter and Ernst Fehr. Fairness in the labour market: A survey of experimental results. In *Surveys in Experimental Economics: Bargaining, Cooperation and Election Stock Markets*, pages 95–132. Springer, 2002.

[95] Florent Garcin, Boi Faltings, Olivier Donatsch, Ayar Alazzawi, Christophe Bruttin, and Amr Huber. Offline and online evaluation of news recommender systems at swissinfo.ch. In *Proceedings of the 8th ACM Conference on Recommender systems*, pages 169–176, 2014.

[96] Mouzhi Ge, Carla Delgado-Battenfeld, and Dietmar Jannach. Beyond accuracy: Evaluating recommender systems by coverage and serendipity. In *Proceedings of the Fourth ACM Conference on Recommender Systems*, pages 257–260, 2010.

[97] Yingqiang Ge, Shuchang Liu, Zuohui Fu, Juntao Tan, Zelong Li, Shuyuan Xu, Yunqi Li, Yikun Xian, and Yongfeng Zhang. A survey on trustworthy recommender systems. *arXiv preprint arXiv:2207.12515*, 2022.

[98] Thomas George and Srujana Merugu. A scalable collaborative filtering framework based on co-clustering. In *Fifth IEEE International Conference on Data Mining (ICDM'05)*, pages 4–pp. IEEE, 2005.

[99] Ken Goldberg, Theresa Roeder, Dhruv Gupta, and Chris Perkins. Eigentaste: A constant time collaborative filtering algorithm. *Information Retrieval*, 4:133–151, 2001.

[100] Gustavo González, Beatriz López, and Josep Lluís de la Rosa. The emotional factor: An innovative approach to user modelling for recommender systems. In *Workshop on Recommendation and Personalization in e-Commerce*, pages 90–99, 2002.

[101] Irwin Greenberg. An analysis of the eeocc "four-fifths" rule. *Management Science*, 25(8):762–769, 1979.

[102] Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 855–864, 2016.

[103] Guibing Guo, Jie Zhang, Daniel Thalmann, and Neil Yorke-Smith. Etaf: An extended trust antecedents framework for trust prediction. In *2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)*, pages 540–547. IEEE, 2014.

[104] Guibing Guo, Jie Zhang, and Neil Yorke-Smith. Trustsvd: Collaborative filtering with both the explicit and implicit influence of user trust and of item ratings. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 29, 1, 2015.

[105] Guibing Guo, Jie Zhang, and Neil Yorke-Smith. A novel recommendation model regularized with user trust and item ratings. *IEEE Transactions on Knowledge and Data Engineering*, 28(7):1607–1620, 2016.

[106] Sara Hajian and Josep Domingo-Ferrer. A methodology for direct and indirect discrimination prevention in data mining. *IEEE Transactions on Knowledge and Data Engineering*, 25(7):1445–1459, 2012.

[107] Moritz Hardt, Eric Price, and Nati Srebro. Equality of opportunity in supervised learning. *Advances in Neural Information Processing Systems*, 29, 2016.

[108] F Maxwell Harper and Joseph A Konstan. The movielens datasets: History and context. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 5(4):1–19, 2015.

[109] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. Neural collaborative filtering. In *Proceedings of the 26th International Conference on World Wide Web*, pages 173–182, 2017.

[110] Denis Helic. Regular equivalence in informed network search. In *37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pages 1088–1093. IEEE, 2014.

[111] Jonathan L Herlocker, Joseph A Konstan, Al Borchers, and John Riedl. An algorithmic framework for performing collaborative filtering. In *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 230–237, 1999.

[112] Jonathan L Herlocker, Joseph A Konstan, Loren G Terveen, and John T Riedl. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)*, 22(1):5–53, 2004.

[113] Balázs Hidasi and Alexandros Karatzoglou. Recurrent neural networks with top-k gains for session-based recommendations. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 843–852, 2018.

[114] Douglas L Hintzman. Minerva 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*, 16(2):96–101, 1984.

[115] Douglas L Hintzman. Schema abstraction in a multiple-trace memory model. *Psychological Review*, 93(4):411, 1986.

[116] Douglas L Hintzman. Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, 95(4):528, 1988.

[117] Frank Hopfgartner, Torben Brodt, Jonas Seiler, Benjamin Kille, Andreas Lommatzsch, Martha Larson, Roberto Turrin, and András Serény. Benchmarking news recommendations: The clef newsreel use case. In *ACM SIGIR Forum*, volume 49, 2, pages 129–136. ACM New York, NY, USA, 2016.

[118] Longke Hu, Aixin Sun, and Yong Liu. Your neighbors affect your ratings: On geographical neighborhood influence to rating prediction. In *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*, pages 345–354, 2014.

[119] Yifan Hu, Yehuda Koren, and Chris Volinsky. Collaborative filtering for implicit feedback datasets. In *2008 Eighth IEEE International Conference on Data Mining*, pages 263–272. IEEE, 2008.

[120] Dietmar Jannach and Gediminas Adomavicius. Recommendations with a purpose. In *Proceedings of the 10th ACM Conference on Recommender Systems*, pages 7–10, 2016.

[121] Dietmar Jannach and Christine Bauer. Escaping the mcnamara fallacy: Towards more impactful recommender systems research. *AI Magazine*, 41(4):79–95, 2020.

[122] Dietmar Jannach and Michael Jugovac. Measuring the business value of recommender systems. *ACM Transactions on Management Information Systems (TMIS)*, 10(4):1–23, 2019.

[123] Dietmar Jannach, Lukas Lerche, and Markus Zanker. Recommending based on implicit feedback. In *Social Information Access: Systems and Technologies*, pages 510–569. Springer, 2018.

[124] Dietmar Jannach and Malte Ludewig. When recurrent neural networks meet the neighborhood for session-based recommendation. In *Proceedings of the Eleventh ACM Conference on Recommender Systems*, pages 306–310, 2017.

[125] Dietmar Jannach, Massimo Quadrana, and Paolo Cremonesi. Session-based recommender systems. In *Recommender Systems Handbook*, pages 301–334. Springer, 2022.

[126] Dietmar Jannach, Paul Resnick, Alexander Tuzhilin, and Markus Zanker. Recommender systems — beyond matrix completion. *Communications of the ACM*, 59(11):94–102, 2016.

[127] Dietmar Jannach and Markus Zanker. Impact and value of recommender systems. *Recommender Systems Handbook*, 2022.

[128] Dietmar Jannach, Markus Zanker, Alexander Felfernig, and Gerhard Friedrich. *Recommender systems: An introduction*. Cambridge University Press, 2010.

[129] Robert Jäschke, Leandro Marinho, Andreas Hotho, Lars Schmidt-Thieme, and Gerd Stumme. Tag recommendations in folksonomies. In *11th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD 2007)*, pages 506–514. Springer, 2007.

[130] Arjan JP Jeckmans, Michael Beye, Zekeriya Erkin, Pieter Hartel, Reginald L Lagendijk, and Qiang Tang. Privacy in recommender systems. In *Social Media Retrieval*, pages 263–281. Springer, 2013.

[131] Yucheng Jin, Nava Tintarev, and Katrien Verbert. Effects of personal characteristics on music recommender systems with different levels of controllability. In *Proceedings of the 12th ACM Conference on Recommender Systems*, pages 13–21, 2018.

[132] Ion Juvina, Othalia Larue, and Alexander Hough. Modeling valuation and core affect in a cognitive architecture: The impact of valence and arousal on memory and decision-making. *Cognitive Systems Research*, 48:4–24, 2018.

[133] Wang-Cheng Kang and Julian McAuley. Self-attentive sequential recommendation. In *2018 IEEE International Conference on Data Mining (ICDM)*, pages 197–206. IEEE, 2018.

[134] Krishnaram Kenthapadi, Benjamin Le, and Ganesh Venkataraman. Personalized job recommendation system at linkedin: Practical challenges and lessons learned. In *Proceedings of the Eleventh ACM conference on Recommender Systems*, pages 346–347, 2017.

[135] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *Stat*, 1050:1, 2014.

[136] Anastasiia Klimashevskaia, Mehdi Elahi, Dietmar Jannach, Christoph Trattner, and Lars Skjærven. Mitigating popularity bias in recommendation: Potential and limits of calibration approaches. In *International Workshop on Algorithmic Bias in Search and Recommendation*, pages 82–90. Springer, 2022.

[137] Anastasiia Klimashevskaia, Dietmar Jannach, Mehdi Elahi, and Christoph Trattner. A survey on popularity bias in recommender systems. *arXiv preprint arXiv:2308.01118*, 2023.

[138] Bart P Knijnenburg and Alfred Kobsa. Making decisions about privacy: Information disclosure in context-aware recommender systems. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 3(3):1–23, 2013.

[139] Bart P Knijnenburg, Martijn C Willemsen, Zeno Gantner, Hakan Soncu, and Chris Newell. Explaining the user experience of recommender systems. *User Modeling and User-Adapted Interaction*, 22:441–504, 2012.

[140] Yehuda Koren, Robert Bell, and Chris Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009.

[141] Pigi Kouki, Shobeir Fakhraei, James Foulds, Magdalini Eirinaki, and Lise Getoor. Hyper: A flexible and extensible probabilistic framework for hybrid recommender systems. In *Proceedings of the 9th ACM Conference on Recommender Systems*, pages 99–106, 2015.

[142] Dominik Kowald. Lfm user groups, zenodo dataset, doi: https://doi.org/10.5281/zenodo.3475975, 2019.

[143] Dominik Kowald. Datasets to evaluate accuracy, miscalibration and popularity lift in recommendations, zenodo dataset, doi: https://doi.org/10.5281/zenodo.7428435, 2022.

[144] Dominik Kowald. Fair recsys datasets, zenodo dataset, doi: https://doi.org/10.5281/zenodo.6123879, 2022.

[145] Dominik Kowald, Sebastian Dennerlein, Dieter Theiler, Simon Walk, and Christoph Trattner. The social semantic server: A framework to provide services on social semantic network data. In *9th International Conference on Semantic Systems, I-SEMANTICS 2013*, pages 50–54. CEUR, 2013.

[146] Dominik Kowald, Simone Kopeinik, and Elisabeth Lex. The tagrec framework as a toolkit for the development of tag-based recommender systems. In *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization*, pages 23–28, 2017.

[147] Dominik Kowald and Emanuel Lacic. Popularity bias in collaborative filtering-based multimedia recommender systems. In *Advances in Bias and Fairness in Information Retrieval, BIAS 2022*, pages 1–11. Springer, 2022.

[148] Dominik Kowald, Emanuel Lacic, and Christoph Trattner. Tagrec: Towards a standardized tag recommender benchmarking framework. In *Proceedings of the 25th ACM Conference on Hypertext and Social Media*, pages 305–307, 2014.

[149] Dominik Kowald and Elisabeth Lex. Evaluating tag recommender algorithms in real-world folksonomies: A comparative study. In *Proceedings of the 9th ACM Conference on Recommender Systems*, pages 265–268, 2015.

[150] Dominik Kowald and Elisabeth Lex. The influence of frequency, recency and semantic context on the reuse of tags in social tagging systems. In *Proceedings of the 27th ACM Conference on Hypertext and Social Media*, pages 237–242, 2016.

[151] Dominik Kowald*, Elisabeth Lex*, and Markus Schedl. Modeling artist preferences for personalized music recommendations. In *Late-Breaking-Results of the 20th annual conference of the International Society for Music Information Retrieval*, ISMIR '19, 2019, *equal contribution.

[152] Dominik Kowald*, Elisabeth Lex*, and Markus Schedl. Utilizing human memory processes to model genre preferences for personalized music recommendations. In *4th Workshop on Transparency and Explainability in Adaptive Systems through User Modeling Grounded in Psychological Theory*. Association of Computing Machinery, 2020, *equal contribution.

[153] Dominik Kowald*, Gregor Mayr*, Markus Schedl, and Elisabeth Lex. A study on accuracy, miscalibration, and popularity bias in recommendations. In *Advances in Bias and Fairness in Information Retrieval, BIAS 2023*, pages 1–16. Springer, 2023, *equal contribution.

[154] Dominik Kowald, Peter Muellner, Eva Zangerle, Christine Bauer, Markus Schedl, and Elisabeth Lex. Support the underground: Characteristics of beyond-mainstream music listeners. *EPJ Data Science*, 10(1):1–26, 2021.

[155] Dominik Kowald, Subhash Chandra Pujari, and Elisabeth Lex. Temporal effects on hashtag reuse in twitter: A cognitive-inspired hashtag recommendation approach. In *Proceedings of the 26th International Conference on World Wide Web*, pages 1401–1410, 2017.

[156] Dominik Kowald, Markus Reiter-Haas, Simone Kopeinik, Markus Schedl, and Elisabeth Lex. Transparent music preference modeling and recommendation with a model of human memory theory. In *A Human-Centered Perspective of Intelligent Personalized Environments and Systems*, pages 113–136. Springer, 2024.

[157] Dominik Kowald, Markus Schedl, and Elisabeth Lex. The unfairness of popularity bias in music recommendation: A reproducibility study. In *Advances in Information Retrieval: 42nd European Conference on IR Research, ECIR 2020*, pages 35–42. Springer, 2020.

[158] Dominik Kowald, Paul Seitlinger, Tobias Ley, and Elisabeth Lex. The impact of semantic context cues on the user acceptance of tag recommendations: An online study. In *Companion Proceedings of the Web Conference 2018*, pages 1–2, 2018.

[159] Dominik Kowald, Deqing Yang, and Emanuel Lacic. Editorial: Reviews in recommender systems. *Frontiers in Big Data*, 6, 2024.

[160] Mark A Kramer. Nonlinear principal component analysis using autoassociative neural networks. *AIChE Journal*, 37(2):233–243, 1991.

[161] Matt J Kusner, Joshua Loftus, Chris Russell, and Ricardo Silva. Counterfactual fairness. *Advances in Neural Information Processing Systems*, 30, 2017.

[162] Florian Königsdorfer, Armin Haberl, Dominik Kowald, Tony Ross-Hellauer, and Stefan Thalmann. Black box or open science? A study on reproducibility in AI development papers. In *Proceedings of The Hawaii International Conference on System Sciences*, HICSS' 24, 2024.

[163] Emanuel Lacic, Leon Fadljevic, Franz Weissenboeck, Stefanie Lindstaedt, and Dominik Kowald. What drives readership? an online study on user interface types and popularity bias mitigation in news article recommendations. In *European Conference on Information Retrieval*, pages 172–179. Springer, 2022.

[164] Emanuel Lacic, Dominik Kowald, Lukas Eberhard, Christoph Trattner, Denis Parra, and Leandro Balby Marinho. Utilizing online social network and location-based data to recommend products and categories in online marketplaces. In *Mining, Modeling, and Recommending'Things' in Social*

Media: 4th International Workshops, MUSE 2013, Prague, Czech Republic, September 23, 2013, and MSM 2013, Paris, France, May 1, 2013, Revised Selected Papers, pages 96–115. Springer, 2015.

[165] Emanuel Lacic, Dominik Kowald, Markus Reiter-Haas, Valentin Slawicek, and Elisabeth Lex. Beyond accuracy optimization: On the value of item embeddings for student job recommendations. In *International Workshop on Multi-dimensional Information Fusion for User Modeling and Personalization (IFUP'2018) co-located with the 11th ACM International Conference on Web Search and Data Mining (WSDM'2018)*, 2017.

[166] Emanuel Lacic, Dominik Kowald, Matthias Traub, Granit Luzhnica, Jörg Peter Simon, and Elisabeth Lex. Tackling cold-start users in recommender systems with indoor positioning systems. In *Proceedings of the 9th ACM Conference on Recommender Systems*. ACM, 2015.

[167] Emanuel Lacic, Markus Reiter-Haas, Dominik Kowald, Manoj Reddy Dareddy, Junghoo Cho, and Elisabeth Lex. Using autoencoders for session-based job recommendations. *User Modeling and User-Adapted Interaction*, 30:617–658, 2020.

[168] Anja Lambrecht and Catherine Tucker. Algorithmic bias? an empirical study of apparent gender-based discrimination in the display of stem career ads. *Management Science*, 65(7):2966–2981, 2019.

[169] Martha Larson, Allesandro Zito, Babak Loni, Paolo Cremonesi, et al. Towards minimal cecessary data: The case for analyzing training data requirements of recommender algorithms. In *Proceedings of the 2017 FATREC Workshop on Responsible Recommendation*, pages 1–6, 2017.

[170] Angela Y Lee. Effects of implicit memory on memory-based versus stimulus-based brand choice. *Journal of Marketing Research*, 39(4):440–454, 2002.

[171] Oleg Lesota, Alessandro Melchiorre, Navid Rekabsaz, Stefan Brandl, Dominik Kowald, Elisabeth Lex, and Markus Schedl. Analyzing item popularity bias of music recommender systems: are different genders equally affected? In *Proceedings of the 15th ACM Conference on Recommender Systems*, pages 601–606, 2021.

[172] Elisabeth Lex and Dominik Kowald. The impact of time on hashtag reuse in twitter: A cognitive-inspired hashtag recommendation approach. In *Proceedings of the 49th GI Annual Conference*, INFORMATIK '19, 2019.

[173] Elisabeth Lex*, Dominik Kowald*, and Markus Schedl. Modeling popularity and temporal drift of music genre preferences. *Transactions of the International Society for Music Information Retrieval*, 3(1), 2020, *equal contribution.

[174] Elisabeth Lex, Dominik Kowald, Paul Seitlinger, Thi Ngoc Trang Tran, Alexander Felfernig, Markus Schedl, et al. Psychology-informed recommender systems. *Foundations and Trends in Information Retrieval*, 15(2):134–242, 2021.

[175] Zihao Li, Xianzhi Wang, Chao Yang, Lina Yao, Julian McAuley, and Guandong Xu. Exploiting explicit and implicit item relationships for session-based recommendation. In *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, pages 553–561, 2023.

[176] Dawen Liang, Rahul G Krishnan, Matthew D Hoffman, and Tony Jebara. Variational autoencoders for collaborative filtering. In *Proceedings of the 2018 World Wide Web Conference*, pages 689–698, 2018.

[177] Kun Lin, Nasim Sonboli, Bamshad Mobasher, and Robin Burke. Calibration in collaborative filtering recommender systems: A user-centered analysis. In *Proceedings of the 31st ACM Conference on Hypertext and Social Media*, HT '20, page 197–206, New York, NY, USA, 2020. Association for Computing Machinery.

[178] Yujie Lin, Pengjie Ren, Zhumin Chen, Zhaochun Ren, Dongxiao Yu, Jun Ma, Maarten de Rijke, and Xiuzhen Cheng. Meta matrix factorization for federated rating predictions. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 981–990, 2020.

[179] Greg Linden, Brent Smith, and Jeremy York. Amazon. com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing*, 7(1):76–80, 2003.

[180] Kun Liu and Evimaria Terzi. A framework for computing the privacy scores of users in online social networks. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 5(1):1–30, 2010.

[181] Pasquale Lops, Marco De Gemmis, and Giovanni Semeraro. Content-based recommender systems: State of the art and trends. *Recommender Systems Handbook*, pages 73–105, 2010.

[182] Malte Ludewig and Dietmar Jannach. Evaluation of session-based recommendation algorithms. *User Modeling and User-Adapted Interaction*, 28:331–390, 2018.

[183] Xin Luo, Mengchu Zhou, Yunni Xia, and Qingsheng Zhu. An efficient non-negative matrix-factorization-based approach to collaborative filtering for recommender systems. *IEEE Transactions on Industrial Informatics*, 10(2):1273–1284, 2014.

[184] Masoud Mansoury, Himan Abdollahpouri, Mykola Pechenizkiy, Bamshad Mobasher, and Robin Burke. Feedback loop and bias amplification in recommender systems. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, pages 2145–2148, 2020.

[185] Leandro Balby Marinho and Lars Schmidt-Thieme. Collaborative tag recommendations. In *Data Analysis, Machine Learning and Applications: Proceedings of the 31st Annual Conference of the Gesellschaft für Klassifikation eV*, pages 533–540. Springer, 2008.

[186] Paolo Massa and Paolo Avesani. Trust-aware recommender systems. In *Proceedings of the 2007 ACM Conference on Recommender Systems*, pages 17–24, 2007.

[187] Judith Masthoff. Group modeling: Selecting a sequence of television items to suit a group of viewers. *Personalized Digital Television: Targeting Programs to Individual Viewers*, pages 93–141, 2004.

[188] Judith Masthoff. The pursuit of satisfaction: Affective state in group recommender systems. In *International Conference on User Modeling*, pages 297–306. Springer, 2005.

[189] Judith Masthoff. Group recommender systems: Combining individual models. In *Recommender Systems Handbook*, pages 677–702. Springer, 2010.

[190] Judith Masthoff and Amra Delic. Group recommender systems: Beyond preference aggregation. *Recommender Systems Handbook*, page 381, 2022.

[191] Judith Masthoff and Albert Gatt. In pursuit of satisfaction and the prevention of embarrassment: Affective state in group recommender systems. *User Modeling and User-Adapted Interaction*, 16:281–319, 2006.

[192] Sean M McNee, John Riedl, and Joseph A Konstan. Being accurate is not enough: How accuracy metrics have hurt recommender systems. In *CHI'06 Extended Abstracts on Human Factors in Computing Systems*, pages 1097–1101, 2006.

[193] Frank McSherry and Ilya Mironov. Differentially private recommender systems: Building privacy into the netflix prize contenders. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 627–636, 2009.

[194] AKM Nuhil Mehdy, Michael D Ekstrand, Bart P Knijnenburg, and Hoda Mehrpouyan. Privacy as a planned behavior: Effects of situational factors on privacy perceptions and plans. In *Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*, pages 169–178, 2021.

[195] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR)*, 54(6):1–35, 2021.

[196] Alessandro B Melchiorre, Navid Rekabsaz, Emilia Parada-Cabaleiro, Stefan Brandl, Oleg Lesota, and Markus Schedl. Investigating gender fairness of recommendation algorithms in the music domain. *Information Processing & Management*, 58(5):102666, 2021.

[197] Alessandro B Melchiorre, Eva Zangerle, and Markus Schedl. Personality bias of music recommendation algorithms. In *Proceedings of the 14th ACM Conference on Recommender Systems*, pages 533–538, 2020.

[198] Tim Miller. Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267:1–38, 2019.

[199] Nuno Moniz, Luís Torgo, Magdalini Eirinaki, and Paula Branco. A framework for recommendation of highly popular news lacking social feedback. *New Generation Computing*, 35:417–450, 2017.

[200] Marta Moscati, Christian Wallmann, Markus Reiter-Haas, Dominik Kowald, Elisabeth Lex, and Markus Schedl. Files for integrating the act-r framework with collaborative filtering for explainable sequential music recommendation, zenodo dataset, doi: https://doi.org/10.5281/zenodo.7923580, 2023.

[201] Marta Moscati, Christian Wallmann, Markus Reiter-Haas, Dominik Kowald, Elisabeth Lex, and Markus Schedl. Integrating the act-r framework with collaborative filtering for explainable sequential music recommendation. In *Proceedings of the 17th ACM Conference on Recommender Systems*, 2023.

[202] Peter Muellner, Dominik Kowald, and Elisabeth Lex. Robustness of meta matrix factorization against strict privacy constraints. In *Advances in Information Retrieval: 43rd European Conference on IR Research, ECIR 2021*, pages 107–119. Springer, 2021.

[203] Peter Muellner, Stefan Schmerda, Dieter Theiler, Stefanie Lindstaedt, and Dominik Kowald. Towards employing recommender systems for supporting data and algorithm sharing. In *DataEconomy Workshop co-located with the 18th International Conference on emerging Networking EXperiments and Technologies*, CoNext '22, 2022.

[204] Peter Müllner, Elisabeth Lex, and Dominik Kowald. Position paper on simulating privacy dynamics in recommender systems. In *SimuRec Workshops co-located with the 15th ACM Conference on Recommender Systems: RECSYS 2021*, 2021.

[205] Peter Müllner, Elisabeth Lex, Markus Schedl, and Dominik Kowald. Differential privacy in collaborative filtering recommender systems: A review. *Frontiers in Big Data*, 6, 2023.

[206] Peter Müllner, Elisabeth Lex, Markus Schedl, and Dominik Kowald. Reuseknn: Neighborhood reuse for differentially private knn-based recommendations. *ACM Trans. Intell. Syst. Technol.*, 14(5), 2023.

[207] Peter Müllner, Elisabeth Lex, Markus Schedl, and Dominik Kowald. The impact of differential privacy on recommendation accuracy and popularity bias. In *European Conference on Information Retrieval*, pages 466–482. Springer, 2024.

[208] Peter Müllner, Dominik Kowald, and Elisabeth Lex. User groups for robustness of meta matrix factorization against decreasing privacy budgets, zenodo dataset, doi: https://doi.org/10.5281/zenodo.4031011, 2020.

[209] Peter Müllner, Dominik Kowald, Markus Schedl, Christine Bauer, Evarle Zange, and Elisabeth Lex. Lfm-beyms, zenodo dataset, doi: https://doi.org/10.5281/zenodo.3784765, 2020.

[210] Athanasios N Nikolakopoulos, Xia Ning, Christian Desrosiers, and George Karypis. Trust your neighbors: A comprehensive survey of neighborhood-based methods for recommender systems. *Recommender Systems Handbook*, pages 39–89, 2021.

[211] Xia Ning and George Karypis. Slim: Sparse linear methods for top-n recommender systems. In *2011 IEEE 11th International Conference on Data Mining*, pages 497–506. IEEE, 2011.

[212] Eirini Ntoutsi, Pavlos Fafalios, Ujwal Gadiraju, Vasileios Iosifidis, Wolfgang Nejdl, Maria-Esther Vidal, Salvatore Ruggieri, Franco Turini, Symeon Papadopoulos, Emmanouil Krasanakis, et al. Bias in data-driven artificial intelligence systems — an introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3):e1356, 2020.

[213] Ingrid Nunes and Dietmar Jannach. A systematic review and taxonomy of explanations in decision support and recommender systems. *User Modeling and User-Adapted Interaction*, 27:393–444, 2017.

[214] John O'Donovan and Barry Smyth. Trust in recommender systems. In *Proceedings of the 10th International Conference on Intelligent User Interfaces*, pages 167–174, 2005.

[215] Özlem Özgöbek, Jon Atle Gulla, and R Cenk Erdur. A survey on challenges and methods in news recommendation. In *International Conference on Web Information Systems and Technologies*, volume 2, pages 278–285. SCITEPRESS, 2014.

[216] Alexandros Paramythis, Stephan Weibelzahl, and Judith Masthoff. Layered evaluation of interactive adaptive systems: Framework and formative methods. *User Modeling and User-Adapted Interaction*, 20:383–453, 2010.

[217] C Whan Park and V Parker Lessig. Familiarity and its impact on consumer decision biases and heuristics. *Journal of Consumer Research*, 8(2):223–230, 1981.

[218] Yoon-Joo Park and Alexander Tuzhilin. The long tail of recommender systems and how to leverage it. In *Proceedings of the 2008 ACM Conference on Recommender Systems*, pages 11–18, 2008.

[219] Denis Parra, Peter Brusilovsky, and Christoph Trattner. See what you want to see: Visual user-driven approach for hybrid recommendation. In

Proceedings of the 19th International Conference on Intelligent User Interfaces, pages 235–240, 2014.

[220] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 701–710, 2014.

[221] Pearl Pu, Li Chen, and Rong Hu. A user-centric evaluation framework for recommender systems. In *Proceedings of the Fifth ACM Conference on Recommender Systems*, pages 157–164, 2011.

[222] Massimo Quadrana, Paolo Cremonesi, and Dietmar Jannach. Sequence-aware recommender systems. *ACM Computing Surveys (CSUR)*, 51(4):1–36, 2018.

[223] Massimo Quadrana, Alexandros Karatzoglou, Balázs Hidasi, and Paolo Cremonesi. Personalizing session-based recommendations with hierarchical recurrent neural networks. In *Proceedings of the Eleventh ACM Conference on Recommender Systems*, pages 130–137, 2017.

[224] Naren Ramakrishnan, Benjamin J Keller, Batul J Mirza, Ananth Y Grama, and George Karypis. When being weak is brave: Privacy in recommender systems. *IEEE Internet Computing*, pages 54–62, 2001.

[225] Matěj Račinský. Myanimelist dataset, kaggle dataset, doi: https://doi.org/10.34740/kaggle/dsv/45582, 2018.

[226] Markus Reiter-Haas, Emilia Parada-Cabaleiro, Markus Schedl, Elham Motamedi, Marko Tkalcic, and Elisabeth Lex. Predicting music relistening behavior using the act-r framework. In *Proceedings of the 15th ACM Conference on Recommender Systems*, pages 702–707, 2021.

[227] Markus Reiter-Haas, David Wittenbrink, and Emanuel Lacic. On the heterogeneous information needs in the job domain: A unified platform for student career. In *Proceedings of the 14th ACM Conference on Recommender Systems*, pages 573–574, 2020.

[228] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. Bpr: Bayesian personalized ranking from implicit feedback. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, pages 452–461, 2009.

[229] Paul Resnick, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom, and John Riedl. Grouplens: An open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work*, pages 175–186, 1994.

[230] Paul Resnick and Hal R Varian. Recommender systems. *Communications of the ACM*, 40(3):56–58, 1997.

[231] Francesco Ricci, Lior Rokach, and Bracha Shapira. Introduction to recommender systems handbook. In *Recommender Systems Handbook*, pages 1–35. Springer, 2010.

[232] Elaine Rich. User modeling via stereotypes. *Cognitive Science*, 3(4):329–354, 1979.

[233] Mark Sanderson et al. Test collection based evaluation of information retrieval systems. *Foundations and Trends in Information Retrieval*, 4(4):247–375, 2010.

[234] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th International Conference on World Wide Web*, pages 285–295, 2001.

[235] Markus Schedl. The lfm-1b dataset for music retrieval and recommendation. In *Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval*, pages 103–110, 2016.

[236] Markus Schedl, Christine Bauer, Wolfgang Reisinger, Dominik Kowald, and Elisabeth Lex. Listener modeling and context-aware music recommendation based on country archetypes. *Frontiers in Artificial Intelligence*, 3, 2021.

[237] Markus Schedl, Stefan Brandl, Oleg Lesota, Emilia Parada-Cabaleiro, David Penz, and Navid Rekabsaz. Lfm-2b: A dataset of enriched music listening events for recommender systems research and fairness analysis. In *Proceedings of the 2022 Conference on Human Information Interaction and Retrieval*, pages 337–341, 2022.

[238] Markus Schedl and Bruce Ferwerda. Large-scale analysis of group-specific music genre taste from collaborative tags. In *2017 IEEE International Symposium on Multimedia (ISM)*, pages 479–482. IEEE, 2017.

[239] Markus Schedl, Peter Knees, Brian McFee, Dmitry Bogdanov, and Marius Kaminskas. Music recommender systems. *Recommender Systems Handbook*, pages 453–492, 2015.

[240] Markus Schedl, Hamed Zamani, Ching-Wei Chen, Yashar Deldjoo, and Mehdi Elahi. Current challenges and visions in music recommender systems research. *International Journal of Multimedia Information Retrieval*, 7:95–116, 2018.

[241] Andrew I Schein, Alexandrin Popescul, Lyle H Ungar, and David M Pennock. Methods and metrics for cold-start recommendations. In *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 253–260, 2002.

[242] Sebastian Scher, Bernhard Geiger, Simone Kopeinik, Andreas Trügler, and Dominik Kowald. A conceptual model for leaving the data-centric approach in machine learning. *arXiv preprint arXiv:2302.03361*, 2023.

[243] Sebastian Scher, Simone Kopeinik, Andreas Trügler, and Dominik Kowald. Modelling the long-term fairness dynamics of data-driven targeted help on job seekers. *Scientific Reports*, 13(1):1727, 2023.

[244] Paul Seitlinger, Dominik Kowald, Simone Kopeinik, Ilire Hasani-Mavriqi, Elisabeth Lex, and Tobias Ley. Attention please! a hybrid resource recommender mimicking attention-interpretation dynamics. In *Proceedings of the 24th International Conference on World Wide Web*, pages 339–345, 2015.

[245] Paul Seitlinger, Dominik Kowald, Christoph Trattner, and Tobias Ley. Recommending tags with a model of human categorization. In *Proceedings of the 22nd ACM International Conference on Information & Knowledge Management*, pages 2381–2386, 2013.

[246] Paul Seitlinger, Tobias Ley, Dominik Kowald, Dieter Theiler, Ilire Hasani-Mavriqi, Sebastian Dennerlein, Elisabeth Lex, and Dietrich Albert. Balancing the fluency-consistency tradeoff in collaborative information search with a recommender approach. *International Journal of Human–Computer Interaction*, 34(6):557–575, 2018.

[247] Harald Semmelrock, Simone Kopeinik, Dieter Theiler, Tony Ross-Hellauer, and Dominik Kowald. Reproducibility in machine learning-driven research. *arXiv preprint arXiv:2307.10320*, 2023.

[248] Yue Shi, Martha Larson, and Alan Hanjalic. List-wise learning to rank with matrix factorization for collaborative filtering. In *Proceedings of the Fourth ACM Conference on Recommender Systems*, pages 269–272, 2010.

[249] Yue Shi, Martha Larson, and Alan Hanjalic. Collaborative filtering beyond the user-item matrix: A survey of the state of the art and future challenges. *ACM Computing Surveys (CSUR)*, 47(1):1–45, 2014.

[250] Rashmi Sinha and Kirsten Swearingen. The role of transparency in recommender systems. In *CHI'02 Extended Abstracts on Human Factors in Computing Systems*, pages 830–831, 2002.

[251] Manel Slokom, Alan Hanjalic, and Martha Larson. Towards user-oriented privacy for recommender system data: A personalization-based approach to gender obfuscation for user profiles. *Information Processing & Management*, 58(6):102722, 2021.

[252] Jessie J Smith, Anas Buhayh, Anushka Kathait, Pradeep Ragothaman, Nicholas Mattei, Robin Burke, and Amy Voida. The many faces of fairness: Exploring the institutional logics of multistakeholder microlending recommendation. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, pages 1652–1663, 2023.

[253] Jessie J Smith, Aishwarya Satwani, Robin Burke, and Casey Fiesler. Recommend me? designing fairness metrics with providers. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency*, pages 2389–2399, 2024.

[254] Nasim Sonboli, Robin Burke, Michael Ekstrand, and Rishabh Mehrotra. The multisided complexity of fairness in recommender systems. *AI magazine*, 43(2):164–176, 2022.

[255] Harald Steck. Calibrated recommendations. In *Proceedings of the 12th ACM Conference on Recommender Systems*, RecSys '18, page 154–162, New York, NY, USA, 2018. Association for Computing Machinery.

[256] Martin Stettinger, Alexander Felfernig, Gerhard Leitner, and Stefan Reiterer. Counteracting anchoring effects in group decision making. In *23rd International Conference on User Modeling, Adaptation and Personalization (UMAP 2015)*, pages 118–130. Springer, 2015.

[257] Martin Stettinger, Alexander Felfernig, Gerhard Leitner, Stefan Reiterer, and Michael Jeran. Counteracting serial position effects in the choicla group decision support environment. In *Proceedings of the 20th International Conference on Intelligent User Interfaces*, pages 148–157, 2015.

[258] Zhu Sun, Jie Yang, Jie Zhang, Alessandro Bozzon, Long-Kai Huang, and Chi Xu. Recurrent knowledge graph embedding for effective recommendation. In *Proceedings of the 12th ACM Conference on Recommender Systems*, pages 297–305, 2018.

[259] Özge Sürer, Robin Burke, and Edward C Malthouse. Multistakeholder recommendation with provider constraints. In *Proceedings of the 12th ACM Conference on Recommender Systems*, pages 54–62, 2018.

[260] Nava Tintarev, Matt Dennis, and Judith Masthoff. Adapting recommendation diversity to openness to experience: a study of human behaviour. In *International Conference on User Modeling, Adaptation, and Personalization*, pages 190–202. Springer, 2013.

[261] Nava Tintarev and Judith Masthoff. A survey of explanations in recommender systems. In *2007 IEEE 23rd International Conference on Data Engineering Workshop*, pages 801–810. IEEE, 2007.

[262] Nava Tintarev and Judith Masthoff. Designing and evaluating explanations for recommender systems. In *Recommender Systems Handbook*, pages 479–510. Springer, 2010.

[263] Nava Tintarev and Judith Masthoff. Explaining recommendations: Design and evaluation. In *Recommender Systems Handbook*, pages 353–382. Springer, 2015.

[264] Marko Tkalcic and Li Chen. Personality and recommender systems. In *Recommender Systems Handbook*, pages 715–739. Springer, 2015.

[265] Thi Ngoc Trang Tran, Alexander Felfernig, and Nava Tintarev. Humanized recommender systems: State-of-the-art and research issues. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 11(2):1–41, 2021.

[266] Christoph Trattner, Dominik Kowald, and Emanuel Lacic. Tagrec: towards a toolkit for reproducible evaluation and development of tag-based recommender algorithms. *ACM SIGWEB Newsletter*, 2015(Winter):1–10, 2015.

[267] Christoph Trattner, Dominik Kowald, Paul Seitlinger, Tobias Ley, Simone Kopeinik, et al. Modeling activation processes in human memory to predict the use of tags in social bookmarking systems. *The Journal of Web Science*, 2(1):1–16, 2016.

[268] Endel Tulving. What is episodic memory? *Current Drections in Psychological Science*, 2(3):67–70, 1993.

[269] Sarah Underwood. Regulation of ai remains elusive, communications of the acm news, url: https://cacm.acm.org/news/248474-regulation-of-ai-remains-elusive/, 2020. Accessed on November 22nd, 2023.

[270] André Calero Valdez and Martina Ziefle. The users' perspective on the privacy-utility trade-offs in health recommender systems. *International Journal of Human-Computer Studies*, 121:108–121, 2019.

[271] Joaquin Vanschoren. Meta-learning. *Automated Machine Learning: Methods, Systems, Challenges*, pages 35–61, 2019.

[272] Saúl Vargas and Pablo Castells. Rank and relevance in novelty and diversity metrics for recommender systems. In *Proceedings of the Fifth ACM Conference on Recommender Systems*, pages 109–116, 2011.

[273] Sahil Verma and Julia Rubin. Fairness definitions explained. In *Proceedings of the International Workshop on Software Fairness*, pages 1–7, 2018.

[274] Pascal Vincent, Hugo Larochelle, Yoshua Bengio, and Pierre-Antoine Manzagol. Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th International Conference on Machine Learning*, pages 1096–1103, 2008.

[275] Hao Wang, Naiyan Wang, and Dit-Yan Yeung. Collaborative deep learning for recommender systems. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1235–1244, 2015.

[276] Yifan Wang, Weizhi Ma, Min Zhang, Yiqun Liu, and Shaoping Ma. A survey on the fairness of recommender systems. *ACM Transactions on Information Systems*, 41(3):1–43, 2023.

[277] Stanley L Warner. Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309):63–69, 1965.

[278] Udi Weinsberg, Smriti Bhagat, Stratis Ioannidis, and Nina Taft. Blurme: Inferring and obfuscating user gender based on ratings. In *Proceedings of the Sixth ACM Conference on Recommender Systems*, pages 195–202, 2012.

[279] Cort J Willmott and Kenji Matsuura. Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Climate Research*, 30(1):79–82, 2005.

[280] Raymond E. Wright. Logistic regression. *Reading and Understanding Multivariate Statistics*, pages 217–244, 1995.

[281] Ke Yang and Julia Stoyanovich. Measuring fairness in ranked outputs. In *Proceedings of the 29th International Conference on Scientific and Statistical Database Management*, pages 1–6, 2017.

[282] Qiang Yang, Yang Liu, Tianjian Chen, and Yongxin Tong. Federated machine learning: concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(2):1–19, 2019.

[283] Eva Zangerle and Christine Bauer. Evaluating recommender systems: Survey and framework. *ACM Computing Surveys*, 55(8):1–38, 2022.

[284] Guijuan Zhang, Yang Liu, and Xiaoning Jin. A survey of autoencoder-based recommender systems. *Frontiers of Computer Science*, 14:430–450, 2020.

[285] Mingwu Zhang, Yu Chen, and Jingqiang Lin. A privacy-preserving optimization of neighbourhood-based recommendation for medical-aided diagnosis and treatment. *IEEE Internet of Things Journal*, 2021.

[286] Minxing Zhang, Zhaochun Ren, Zihan Wang, Pengjie Ren, Zhunmin Chen, Pengfei Hu, and Yang Zhang. Membership inference attacks against recommender systems. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, pages 864–879, 2021.

[287] Shijie Zhang, Wei Yuan, and Hongzhi Yin. Comprehensive privacy analysis on federated recommender system against attribute inference attacks. *IEEE Transactions on Knowledge and Data Engineering*, 2023.

[288] Shuai Zhang, Lina Yao, Aixin Sun, and Yi Tay. Deep learning based recommender system: A survey and new perspectives. *ACM Computing Surveys (CSUR)*, 52(1):1–38, 2019.

[289] Yuan Cao Zhang, Diarmuid Ó Séaghdha, Daniele Quercia, and Tamas Jambor. Auralist: Introducing serendipity into music recommendation. In *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining*, pages 13–22, 2012.

[290] Cai-Nicolas Ziegler, Sean M McNee, Joseph A Konstan, and Georg Lausen. Improving recommendation lists through topic diversification. In *Proceedings of the 14th International Conference on World Wide Web*, pages 22–32, 2005.