Multiperspective and Multidisciplinary Treatment of Fairness in Recommender Systems Research

Markus Schedl
Navid Rekabsaz
markus.schedl@jku.at
navid.rekabsaz@jku.at
Johannes Kepler University (Insitute of Computational Perception) and Linz Institute of Technology (AI Lab)
Linz, Austria

Elisabeth Lex
elisabeth.lex@tugraz.at
Graz University of Technology (Institute of Interactive Systems and Data Science)
Graz, Austria

Tessa Grosz
Elisabeth Greif
tessa.grosz@jku.at
elisabeth.greif@jku.at
Johannes Kepler University (Institute for Legal Gender Studies)
Linz, Austria

ABSTRACT

In the communities of UMAP, RecSys, and similar venues, fairness of recommender systems has primarily been addressed from the perspective of computer science and artificial intelligence, e.g., by devising computational bias and fairness metrics or elaborating debiasing algorithms. In contrast, we advocate taking a multiperspective and multidisciplinary viewpoint to complement this technical perspective. This involves considering the variety of stakeholders in the value chain of recommender systems as well as interweaving expertise from various disciplines, in particular, computer science, law, ethics, sociology, and psychology (e.g., studying discrepancies between computational metrics of bias and fairness and their actual human perception, and considering the legal and regulatory context recommender systems are embedded in).

ACM Reference Format:
Markus Schedl, Navid Rekabsaz, Elisabeth Lex, Tessa Grosz, and Elisabeth Greif. 2022. Multiperspective and Multidisciplinary Treatment of Fairness in Recommender Systems Research. In *Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '22 Adjunct), July 4–7, 2022, Barcelona, Spain.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3511047.3536400

1 INTRODUCTION

Recommender systems (RSs) impact our day-to-day decisions by, e.g., deciding on the content we are exposed to on the web, or what products to buy, thereby narrowing our view of the world. They create for their users personalized recommendation lists of items selected from a commonly very large item catalog [43]. While RSs are incredibly useful tools that guide their users through the massive amounts of digital content nowadays available, several studies, e.g., [16, 27, 30, 31, 34, 42, 54], show that different parts of a RS can cause or amplify manifold biases. Some of these biases can be harmful because they result in unfair or discriminating treatment of certain users or groups of users [17, 19], which often is ethically and legally problematic.

While there are numerous efforts in the research communities behind UMAP, RecSys, SIGIR, and similar venues to identify biases in RSs and ensure RSs treat users fairly, they often view the problem from a technical and quantitative perspective [21]. However, there is no clear consensus in the community about what is considered biased or fair,[1] how to measure these aspects, and what the social and legal implications of a biased result of a RS will be in reality.

In this paper, we advocate to establish a broader view on the topic of fairness in RSs research, including psychological, sociological, legal, and regulatory considerations, thereby promoting the following emerging research topics:

(1) Considering the variety of stakeholders in the value chain of RSs when investigating biases and fairness of respective technological solutions (discussed in Section 2).
(2) Raising awareness of likely discrepancies between computational metrics of bias and fairness and their actual individual and societal perception (Section 3).
(3) Discussing metrics of bias and fairness as well as technical debiasing solution in the context of ethical considerations and legal regulations (Section 4).

1.1 Dimensions of Detrimental Biases

Potentially harmful biases can be categorized according to different taxonomies, e.g., [4]. However, few if any of them are universally agreed upon in the scientific community. Nevertheless, we briefly introduce some of the most important categorization schemes.

In terms of comparing the real world, an ideal world, and their representation in the RS and output behavior of the RS, we can distinguish between *societal bias* and *statistical bias*. The former refers to the discrepancy between the system behavior and what it should be in an ideal world. The latter refers to the discrepancy between the system behavior and what it should be to reflect the real world, e.g., [16, 27, 34].

Biases can further be categorized w.r.t. the stage in the recommendation process at which they occur or intensify, as illustrated in Figure 1. They can be present already in the *data* the algorithms are trained on (cf. statistical bias above). For instance, when the algorithms are provided with an unbalanced dataset w.r.t. certain attributes, e.g., distribution of genders or ethnicities. Furthermore,

[1] RSs create *personalized* recommendation lists, and therefore a certain level of bias towards items that match the user's preferences is expected.

UMAP '22 Adjunct, July 4–7, 2022, Barcelona, Spain

ACM ISBN 978-1-4503-9232-7/22/07.
https://doi.org/10.1145/3511047.3536400

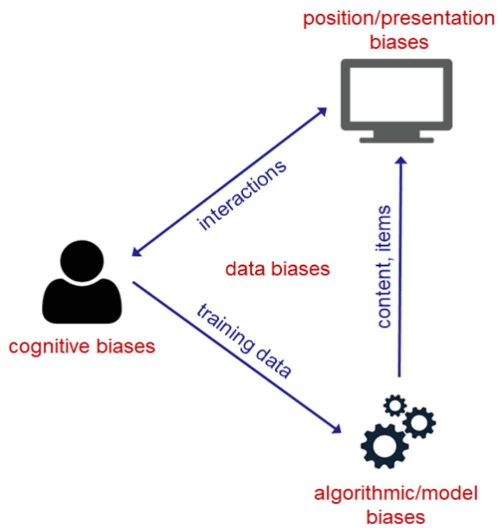

Figure 1: Categories of biases in RSs, according to the stages of the recommendation process.

the RS *algorithm* itself can introduce or amplify biases that are subsequently encoded in the learned *model*. For instance, the models can reinforce stereotypes such as female users predominantly being interested in certain types of (often low-paying) jobs, or emotionally unstable teenagers being fans of metal music. Finally, biases can also result from the interaction between the user and the RS, which includes *presentation* biases of the system and *cognitive* biases of the user. Those are strongly tied, since the way recommended items are presented to the user affects their cognitive processing of the recommendation list. Examples include serial positioning effects [24], i.e., users more likely remember the items at the top and the end of a recommendation list, and decoy effects [46], i.e., items in the recommendation lists that are very inferior to an alternative (target) item can improve the perceived quality of the target item [32].

Harmful biases can also be categorized w.r.t. user or item properties according to which the RS discriminates. The former include *demographic* biases (e.g., providing lower utility for some groups defined by age, gender, or country) [20, 36, 39] and *psychological* biases (treating differently user groups with different psychological traits such as personality) [32, 37, 38]. On the item level, *popularity* bias is the most frequently researched bias, e.g., [2, 20, 22, 29, 31, 33]. It refers to an overrepresentation of already popular items in the recommendation lists of many users.

To quantify detrimental biases and fairness of RSs, various *metrics* have been proposed to measure how much a recommendation model suffers from a particular bias; see the survey by [4] for a recent overview of the most relevant metrics. *Group fairness metrics* such as statistical parity [53] or disparate impact [23] compare the fairness of a recommendation model w.r.t. two or more groups of users. To that end, Yao et al. [51] propose five fairness metrics to quantify whether preferences for one group are consistently overestimated compared to their actual ratings. Abdollahpouri et al. [3] investigate popularity bias and introduce the ΔGAP metric to measure the difference between the average popularity of items in a user group and their recommendation lists. *Individual fairness metrics* measure if similar individuals receive similar treatments [18]. Such metrics consider the distribution of biases and fairness at an individual level using inequality measures, e.g., generalized entropy index [45] or various statistical quantities [31].

1.2 Bias Mitigation Strategies

Approaches to mitigate harmful biases can be categorized according to the stage of the recommendation process at which they are included into pre-processing, in-processing, and post-processing methods [35]. *Pre-processing* approaches focus on preparing the data in such a way that the resulting model and the final recommendations show a lower degree of bias. Examples of such methods are data rebalancing, i.e., upsampling data from the minority group [36], and scraping gender-related words in biographies in content-based job classification/recommendation [15]. *In-processing* methods aim at learning a bias-/fairness-aware model, mainly through adding specific bias/fairness criteria to the optimization procedure. This includes methods such as regularization, where the loss function is extended by a bias correction term, and adversarial learning, where the encoded bias in user-item interaction data is reduced by models' internal representations to become agnostic with regard to a specific protected attribute [8, 25, 42]. Finally, *post-processing* approaches typically adopt re-ranking methods that re-organize a recommendation list created by another recommendation algorithm to enforce a certain bias-reducing ranking of recommended items [6, 52].

2 MULTISTAKEHOLDER PERSPECTIVES

Creating, maintaining, and using a RS service involves different stakeholders, e.g., [1, 5, 47]. The most prominent ones — which are also those targeted most frequently in research on bias mitigation and fairness improvement of RSs— are the *end users* of the system, usually the content *consumers* of products, job announcements, movies, music, etc. Nevertheless, other stakeholders do exist, which are often neglected in research on the topics mentioned above. These include the *creators* of the content that is made available through the RS,[2] the *providers* of the recommendation platform (e.g., Netflix, Spotify, Amazon), and *policymakers and political institutions* (e.g., EU and national legislation).

All of them have different needs, requirements, and goals, which sometimes overlap, but sometimes are in conflict with each other. For instance, consumers are commonly interested in receiving recommendations that match their taste (similarity), are fresh (novelty), or cover a wider range of content (diversity), and may be even unaware of fairness issues in recommender systems [44]. Creators typically aim at increasing their items' exposure, by having them recommended to many consumers that are likely to interact with their items. The goals of the RS providers are strongly tied to their economic success, which may be achieved by realizing an advantage over competitors, often by establishing a unique selling point by including functionality that increases the user experience of the RS. Policymakers, on the other hand, sometimes aim at imposing

[2] Note that, depending on the recommendation domain, the creator of an item may comprise different persons; for instance, in the case of a movie, including directors, screenwriters, producers, actors, and actresses.

certain regulatory limitations, for instance, to require a minimum quota for items by national content creators. A prominent example is France, whose *Conseil supérieur de l'audiovisuel* [3] imposed such quotas, with an obvious direct impact on the bias and fairness of the RS ecosystem.

These examples highlight the need for considering the various stakeholders, with their individual goals, in the discussion of fairness and development of novel RS technology. While first RS approaches targeting multistakeholder fairness have emerged recently, e.g., [7, 48], it is critical to raise awareness among researchers and developers of recommendation algorithms that optimizing fairness for one stakeholder may negatively affect others'.

3 PSYCHOLOGICAL AND SOCIOLOGICAL PERSPECTIVES

We challenge the assumption underlying existing metrics (see Section 1.1) that fairness can be measured in a fully objective way that reflects a common and agreed-upon understanding of what is considered fair or unfair system behavior. Similar to findings of user studies, which have shown that perceived recommendation quality and diversity (among other aspects) differ between users of RSs, depending on demographics and other factors [28, 41], we hypothesize that there also exist *substantial differences between computational bias and fairness metrics and the perception of fairness by individuals or groups of individuals* defined by common traits such as gender, age, ethnicity, cultural background, religion, and beliefs. While this has not yet been investigated in the context of RSs, to the best of our knowledge, psychological studies on verbal gender descriptions revealed differences in the level of bias/stereotypes between demographic groups [26]. Also, neuroscience experiments using transcranial direct current stimulation showed that gender stereotypes differ based on neural activity and cultural background [49].

We, therefore, strongly advocate a more holistic perspective when considering the fairness of RSs, and refrain from considering existing metrics as universal. We believe that much more research is needed to investigate and understand the likely discrepancy between computational notions of bias and fairness in RSs research and their human perception in relation to psychological, sociological, and cultural backgrounds.

4 LEGAL AND REGULATORY PERSPECTIVES

On a legal level, the risks associated with AI applications including RSs are rapidly gaining importance, with the EU playing a special role, as it aims at becoming the global standard setter in the regulation of AI with the Artificial Intelligence Act [14],[4] the Digital Service Act [12],[5] and the Digital Market Act [13].[6] While not yet in force once adopted, these three acts will be directly applicable across the EU. In the meantime, what already applies to RSs, although not specifically tailored to AI, is EU anti-discrimination law [9–11, 40].[7,8,9,10] Despite this high density of AI regulation at the EU level, most of the existing legal literature on bias focuses on US law. We should close this gap.

With regard to bias in RSs, we recognize a significant divergence: While "fairness" is a concept frequently referred to in the RSs literature, it is not a legal concept [50]. Instead of fairness, EU law focuses on the protection of fundamental rights and the principle of non-discrimination. With regard to the latter, EU anti-discrimination law addresses discrimination only in certain areas (mainly employment, access to goods and services, and housing) and only when certain categories (e.g., gender, ethnicity, religious belief) are affected. Whether a biased outcome of a RS is discriminatory in a legal sense will depend on the specific context. Moreover, compliance with fundamental rights or with non-discrimination law is not a question of measurability, but of either-or. By including a multidisciplinary perspective, we should develop a fairness concept for RSs that is both informed by and compliant with EU non-discrimination law.

5 CONCLUSION

We briefly reviewed existing notions of (harmful) biases and of fairness in RSs research, and strategies to mitigate the former while increasing the latter. We propose to take a holistic and multidisciplinary perspective on RSs's biases and fairness. This involves considering the recommendation task as a multistakeholder problem, including consumers, creators, platform providers, political and legal institutions. These actors have different goals, some of which are in agreement, some are contradictory. Determining a tradeoff between optimizing recommendation algorithms for fairness and maximizing accuracy is already not an easy task. It becomes even more complicated when considering the different stakeholders' needs and aims. Besides, we would like to motivate the more technical audience (foremost computer scientists and AI researchers) who design and develop RSs to engage into discussions with experts from psychology, sociology, cultural studies, ethics, and law. Even more, we believe that results of empirical studies and models from these disciplines can be integrated into RSs, as already demonstrated with psychology-informed RSs [32]. Furthermore, we aim at establishing a new notion of fairness, which can be influenced by an individual's traits, rather than being treated as a universal and static aspect of a RS.

We are confident that our statements will contribute to the vibrant research area of *trustworthy artificial intelligence*, in particular research on debiasing and fairness of state-of-the-art RSs.

ACKNOWLEDGMENTS

This work received financial support by the Austrian Science Fund (FWF): P33526 and DFH-23; and by the State of Upper Austria and the Federal Ministry of Education, Science, and Research, through grant LIT-2020-9-SEE-113 and LIT-2021-YOU-215.

[3] https://www.csa.fr

[4] https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021PC0206&from=en

[5] https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020PC0825&from=en

[6] https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020PC0842&from=en

[7] https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32000L0043&from=en

[8] https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32000L0078&from=en

[9] https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32004L0113&from=en

[10] https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32006L0054&from=en

REFERENCES

- [1] Himan Abdollahpouri, Gediminas Adomavicius, Robin Burke, Ido Guy, Dietmar Jannach, Toshihiro Kamishima, Jan Krasnodebski, and Luiz Augusto Pizzato. 2020. Multistakeholder recommendation: Survey and research directions. *User Model. User Adapt. Interact.* 30, 1 (2020), 127–158. https://doi.org/10.1007/s11257-019-09256-1
- [2] Himan Abdollahpouri, Masoud Mansoury, Robin Burke, and Bamshad Mobasher. 2019. The Unfairness of Popularity Bias in Recommendation. In *Proceedings of the Workshop on Recommendation in Multi-stakeholder Environments co-located with the 13th ACM Conference on Recommender Systems (RecSys 2019), Copenhagen, Denmark, September 20, 2019 (CEUR Workshop Proceedings, Vol. 2440)*, Robin Burke, Himan Abdollahpouri, Edward C. Malthouse, K. P. Thai, and Yongfeng Zhang (Eds.). CEUR-WS.org. http://ceur-ws.org/Vol-2440/paper4.pdf
- [3] Himan Abdollahpouri, Masoud Mansoury, Robin Burke, and Bamshad Mobasher. 2019. The Unfairness of Popularity Bias in Recommendation. In *Workshop on Recommendation in Multi-stakeholder Environments (RMSE'19), in conjunction with the 13th ACM Conference on Recommender Systems, RecSys'19*.
- [4] Ashwathy Ashokan and Christian Haas. 2021. Fairness metrics and bias mitigation strategies for rating predictions. *Information Processing & Management* 58, 5 (2021), 102646. https://doi.org/10.1016/j.ipm.2021.102646
- [5] Robin Burke, Himan Abdollahpouri, Edward C. Malthouse, K. P. Thai, and Yongfeng Zhang. 2019. Recommendation in multistakeholder environments. In *Proceedings of the 13th ACM Conference on Recommender Systems, RecSys 2019, Copenhagen, Denmark, September 16-20, 2019*, Toine Bogers, Alan Said, Peter Brusilovsky, and Domonkos Tikk (Eds.). ACM, 566–567. https://doi.org/10.1145/3298689.3346973
- [6] L Elisa Celis, Damian Straszak, and Nisheeth K Vishnoi. 2018. Ranking with Fairness Constraints. In *Proceedings of the 45th International Colloquium on Automata, Languages, and Programming (ICALP)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.
- [7] Harshal A. Chaudhari, Sangdi Lin, and Ondrej Linda. 2020. A General Framework for Fairness in Multistakeholder Recommendations. *CoRR* abs/2009.02423 (2020). arXiv:2009.02423 https://arxiv.org/abs/2009.02423
- [8] Jiawei Chen, Hande Dong, Xiang Wang, Fuli Feng, Meng Wang, and Xiangnan He. 2022 (to appear). Bias and Debias in Recommender System: A Survey and Future Directions. *Transactions on Knowledge and Data Engineering* (2022 (to appear)).
- [9] European Commission. 2000. Directive 2000/43/EC implementing the principle of equal treatment between persons irrespective of racial or ethnic origin.
- [10] European Commission. 2000. Directive 2000/78/EC establishing a general framework for equal treatment in employment and occupation: Framework Directive.
- [11] European Commission. 2004. Directive 2004/113/EC implementing the principle of equal treatment between men and women in the access to and supply of goods and services: Goods and Services Directive.
- [12] European Commission. 2020. Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC: COM(2020) 825 final.
- [13] European Commission. 2020. Proposal for a Regulation of the European Parliament and of the Council on contestable and fair markets in the digital sector (Digital Markets Act): COM(2020) 842 final.
- [14] European Commission. 2021. Proposal for a Regulation of the European Parliament and of the Council laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts: COM(2021) 206 final.
- [15] Maria De-Arteaga, Alexey Romanov, Hanna Wallach, Jennifer Chayes, Christian Borgs, Alexandra Chouldechova, Sahin Geyik, Krishnaram Kenthapadi, and Adam Tauman Kalai. 2019. Bias in bios: A case study of semantic representation bias in a high-stakes setting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*. 120–128.
- [16] Yashar Deldjoo, Vito Walter Anelli, Hamed Zamani, Alejandro Bellogín, and Tommaso Di Noia. 2021. A Flexible Framework for Evaluating User and Item Fairness in Recommender Systems. *User Modeling and User-Adapted Interaction* 31, 3 (2021), 457–511. https://doi.org/10.1007/s11257-020-09285-1
- [17] Tommaso Di Noia, Nava Tintarev, Panagiota Fatourou, and Markus Schedl. 2022. Recommender Systems under European AI Regulations. *Commun. ACM* 65, 4 (mar 2022), 69–73. https://doi.org/10.1145/3512728
- [18] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. 2012. Fairness through awareness. In *Proceedings of the 3rd innovations in theoretical computer science conference*. 214–226.
- [19] Michael D. Ekstrand, Anubrata Das, Robin Burke, and Fernando Diaz. 2022 (to appear). Fairness and Discrimination in Information Access Systems. *Foundations and Trends in Information Retrieval* (2022 (to appear)). arXiv:2105.05779 https://arxiv.org/abs/2105.05779
- [20] Michael D. Ekstrand, Mucun Tian, Ion Madrazo Azpiazu, Jennifer D. Ekstrand, Oghenemaro Anuyah, David McNeill, and Maria Soledad Pera. 2018. 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, FAT 2018, 23-24 February 2018, New York, NY, USA (Proceedings of Machine Learning Research, Vol. 81)*, Sorelle A. Friedler and Christo Wilson (Eds.). PMLR, 172–186. http://proceedings.mlr.press/v81/ekstrand18b.html
- [21] Mehdi Elahi, Himan Abdollahpouri, Masoud Mansoury, and Helma Torkamaan. 2021. Beyond algorithmic fairness in recommender systems. In *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*. 41–46.
- [22] Mehdi Elahi, Danial Khosh Kholgh, Mohammad Sina Kiarostami, Sorush Saghari, Shiva Parsa Rad, and Marko Tkalčič. 2021. Investigating the impact of recommender systems on user-based and item-based popularity bias. *Information Processing & Management* 58, 5 (2021), 102655.
- [23] Michael Feldman, Sorelle A Friedler, John Moeller, Carlos Scheidegger, and Suresh Venkatasubramanian. 2015. Certifying and removing disparate impact. In *proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*. 259–268.
- [24] Alexander Felfernig, Gerhard Friedrich, Bartosz Gula, Martin Hitz, Thomas Kruggel, Gerhard Leitner, Rudolf Melcher, D. Riepan, S. Strauss, Erich Teppan, and Oliver Vitouch. 2007. Persuasive Recommendation: Serial Position Effects in Knowledge-Based Recommender Systems. In *Persuasive Technology, Second International Conference on Persuasive Technology, PERSUASIVE 2007, Palo Alto, CA, USA, April 26-27, 2007, Revised Selected Papers (Lecture Notes in Computer Science, Vol. 4744)*, Yvonne de Kort, Wijnand A. IJsselsteijn, Cees J. H. Midden, Berry Eggen, and B. J. Fogg (Eds.). Springer, 283–294. https://doi.org/10.1007/978-3-540-77006-0_34
- [25] Christian Ganhör, David Penz, Navid Rekabsaz, Oleg Lesota, and Markus Schedl. 2022. Mitigating Consumer Biases in Recommendations with Adversarial Training. In *Proceedings of the 45th International ACM SIGIR conference on research and development in Information Retrieval, SIGIR 2022*. ACM.
- [26] Tanja Hentschel, Madeline E. Heilman, and Claudia V. Peus. 2019. The Multiple Dimensions of Gender Stereotypes: A Current Look at Men's and Women's Characterizations of Others and Themselves. *Frontiers in Psychology* 10 (2019). https://doi.org/10.3389/fpsyg.2019.00011
- [27] Ömer Kirnap, Fernando Diaz, Asia Biega, Michael D. Ekstrand, Ben Carterette, and Emine Yilmaz. 2021. Estimation of Fair Ranking Metrics with Incomplete Judgments. In *WWW '21: The Web Conference 2021, Virtual Event / Ljubljana, Slovenia, April 19-23, 2021*, Jure Leskovec, Marko Grobelnik, Marc Najork, Jie Tang, and Leila Zia (Eds.). ACM / IW3C2, 1065–1075. https://doi.org/10.1145/3442381.3450080
- [28] Bart P. Knijnenburg, Martijn C. Willemsen, Zeno Gantner, Hakan Soncu, and Chris Newell. 2012. Explaining the User Experience of Recommender Systems. *User Modeling and User-Adapted Interaction* 22, 4-5 (2012), 441–504. https://doi.org/10.1007/s11257-011-9118-4
- [29] Dominik Kowald, Markus Schedl, and Elisabeth Lex. 2020. The Unfairness of Popularity Bias in Music Recommendation: A Reproducibility Study. In *European Conference on Information Retrieval*. Springer, 35–42.
- [30] Anja Lambrecht and Catherine Tucker. 2019. Algorithmic Bias? An Empirical Study of Apparent Gender-Based Discrimination in the Display of STEM Career Ads. *Management Science* 65, 7 (2019), 2966–2981. https://doi.org/10.1287/mnsc.2018.3093
- [31] Oleg Lesota, Alessandro B. Melchiorre, Navid Rekabsaz, Stefan Brandl, Dominik Kowald, Elisabeth Lex, and Markus Schedl. 2021. Analyzing Item Popularity Bias of Music Recommender Systems: Are Different Genders Equally Affected?. In *Proceedings of the 15th ACM Conference on Recommender Systems (Late-Breaking Results)*. Amsterdam, the Netherlands.
- [32] Elisabeth Lex, Dominik Kowald, Paul Seitlinger, Thi Ngoc Trang Tran, Alexander Felfernig, and Markus Schedl. 2021. Psychology-informed Recommender Systems. *Foundations and Trends® in Information Retrieval* 15, 2 (2021), 134–242.
- [33] Masoud Mansoury, Himan Abdollahpouri, Mykola Pechenizkiy, Bamshad Mobasher, and Robin Burke. 2020. *FairMatch: A Graph-Based Approach for Improving Aggregate Diversity in Recommender Systems*. Association for Computing Machinery, New York, NY, USA, 154–162. https://doi.org/10.1145/3340631.3394860
- [34] Masoud Mansoury, Bamshad Mobasher, Robin Burke, and Mykola Pechenizkiy. 2019. Bias Disparity in Collaborative Recommendation: Algorithmic Evaluation and Comparison. In *Proceedings of the Workshop on Recommendation in Multi-stakeholder Environments co-located with the 13th ACM Conference on Recommender Systems (RecSys 2019), Copenhagen, Denmark, September 20, 2019 (CEUR Workshop Proceedings, Vol. 2440)*, Robin Burke, Himan Abdollahpouri, Edward C. Malthouse, K. P. Thai, and Yongfeng Zhang (Eds.). CEUR-WS.org. http://ceur-ws.org/Vol-2440/paper6.pdf
- [35] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. 2021. A Survey on Bias and Fairness in Machine Learning. *Comput. Surveys* 54, 6 (2021), 115:1–115:35. https://doi.org/10.1145/3457607
- [36] Alessandro B. Melchiorre, Navid Rekabsaz, Emilia Parada-Cabaleiro, Stefan Brandl, Oleg Lesota, and Markus Schedl. 2021. Investigating Gender Fairness of Recommendation Algorithms in the Music Domain. *Information Processing & Management* 58, 5 (2021), 102666. https://doi.org/10.1016/j.ipm.2021.102666
- [37] Alessandro B. Melchiorre and Markus Schedl. 2020. Personality Correlates of Music Audio Preferences for Modelling Music Listeners. In *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization, UMAP*

2020, Genoa, Italy, July 12-18, 2020, Tsvi Kuflik, Ilaria Torre, Robin Burke, and Cristina Gena (Eds.). ACM, 313–317. https://doi.org/10.1145/3340631.3394874

[38] Alessandro B. Melchiorre, Eva Zangerle, and Markus Schedl. 2020. Personality Bias of Music Recommendation Algorithms. In *Fourteenth ACM Conference on Recommender Systems* (Virtual Event, Brazil) *(RecSys '20)*. Association for Computing Machinery, New York, NY, USA, 533–538. https://doi.org/10.1145/3383313.3412223

[39] Nicola Neophytou, Bhaskar Mitra, and Catherine Stinson. 2021. Revisiting Popularity and Demographic Biases in Recommender Evaluation and Effectiveness. *CoRR* abs/2110.08353 (2021). arXiv:2110.08353 https://arxiv.org/abs/2110.08353

[40] European Parliament and the Council. 2006. Directive 2006/54/EC on the implementation of the principle of equal opportunities and equal treatment of men and women in matters of employment and occupation (recast): Gender Recast Directive.

[41] Pearl Pu, Li Chen, and Rong Hu. 2011. A User-centric Evaluation Framework for Recommender Systems. In *Proceedings of the 2011 ACM Conference on Recommender Systems, RecSys 2011, Chicago, IL, USA, October 23-27, 2011*, Bamshad Mobasher, Robin D. Burke, Dietmar Jannach, and Gediminas Adomavicius (Eds.). ACM, 157–164. https://doi.org/10.1145/2043932.2043962

[42] Navid Rekabsaz, Simone Kopeinik, and Markus Schedl. 2021. Societal Biases in Retrieved Contents: Measurement Framework and Adversarial Mitigation of BERT Rankers. In *SIGIR '21: The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Virtual Event, Canada, July 11-15, 2021*, Fernando Diaz, Chirag Shah, Torsten Suel, Pablo Castells, Rosie Jones, and Tetsuya Sakai (Eds.). ACM, 306–316. https://doi.org/10.1145/3404835.3462949

[43] Francesco Ricci, Lior Rokach, and Bracha Shapira (Eds.). 2015. *Recommender Systems Handbook* (2nd ed.). Springer. https://doi.org/10.1007/978-1-4899-7637-6

[44] Nasim Sonboli, Jessie J Smith, Florencia Cabral Berenfus, Robin Burke, and Casey Fiesler. 2021. Fairness and transparency in recommendation: The users' perspective. In *Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*. 274–279.

[45] Till Speicher, Hoda Heidari, Nina Grgic-Hlaca, Krishna P Gummadi, Adish Singla, Adrian Weller, and Muhammad Bilal Zafar. 2018. A unified approach to quantifying algorithmic unfairness: Measuring individual &group unfairness via inequality indices. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2239–2248.

[46] Erich Christian Teppan and Alexander Felfernig. 2009. Impacts of decoy elements on result set evaluations in knowledge-based recommendation. *Int. J. Adv. Intell. Paradigms* 1, 3 (2009), 358–373. https://doi.org/10.1504/IJAIP.2009.026573

[47] Dandan Wang and Yan Chen. 2020. A Novel Many-Objective Recommendation Algorithm for Multistakeholders. *IEEE Access* 8 (2020), 196482–196499. https://doi.org/10.1109/ACCESS.2020.3034716

[48] Dandan Wang and Yan Chen. 2021. A novel cascade hybrid many-objective recommendation algorithm incorporating multistakeholder concerns. *Inf. Sci.* 577 (2021), 105–127. https://doi.org/10.1016/j.ins.2021.07.005

[49] Siqi Wang, Jinjin Wang, Wenmin Guo, Hang Ye, Xinbo Lu, Jun Luo, and Haoli Zheng. 2019. Gender Difference in Gender Bias: Transcranial Direct Current Stimulation Reduces Male's Gender Stereotypes. *Frontiers in Human Neuroscience* 13 (2019). https://doi.org/10.3389/fnhum.2019.00403

[50] Alice Xiang and Inioluwa Deborah Raji. 2019. On the Legal Compatibility of Fairness Definitions. *CoRR* abs/1912.00761 (2019). arXiv:1912.00761 http://arxiv.org/abs/1912.00761

[51] Sirui Yao and Bert Huang. 2017. Beyond parity: Fairness objectives for collaborative filtering. *Advances in neural information processing systems* 30 (2017).

[52] Meike Zehlike, Francesco Bonchi, Carlos Castillo, Sara Hajian, Mohamed Megahed, and Ricardo Baeza-Yates. 2017. Fa*ir: A fair top-k ranking algorithm. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*. 1569–1578.

[53] Rich Zemel, Yu Wu, Kevin Swersky, Toni Pitassi, and Cynthia Dwork. 2013. Learning fair representations. In *International conference on machine learning*. PMLR, 325–333.

[54] Ziwei Zhu, Jianling Wang, and James Caverlee. 2020. Measuring and Mitigating Item Under-Recommendation Bias in Personalized Ranking Systems. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval, SIGIR 2020, Virtual Event, China, July 25-30, 2020*, Jimmy Huang, Yi Chang, Xueqi Cheng, Jaap Kamps, Vanessa Murdock, Ji-Rong Wen, and Yiqun Liu (Eds.). ACM, 449–458. https://doi.org/10.1145/3397271.3401177