Fairness in Ranking: From Values to Technical Choices and Back

Julia Stoyanovich
stoyanovich@nyu.edu
New York University
New York, NY, USA

Meike Zehlike
meike.zehlike@zalando.de
Zalando
Berlin, Germany

Ke Yang
key@umass.edu
University of Massachusetts Amherst
Amherst, MA, USA

ABSTRACT

In the past few years, there has been much work on incorporating fairness requirements into the design of algorithmic rankers, with contributions from the data management, algorithms, information retrieval, and recommender systems communities. In this tutorial, we give a systematic overview of this work, offering a broad perspective that connects formalizations and algorithmic approaches across subfields.

During the first part of the tutorial, we present a classification framework for fairness-enhancing interventions, along which we will then relate the technical methods. This framework allows us to unify the presentation of mitigation objectives and of algorithmic techniques to help meet those objectives or identify trade-offs. Next, we discuss fairness in score-based ranking and in supervised learning-to-rank. We conclude with recommendations for practitioners, to help them select a fair ranking method based on the requirements of their specific application domain.

CCS CONCEPTS

• **Information systems** → **Data management systems**; • **Social and professional topics** → **Computing / technology policy**.

KEYWORDS

algorithmic fairness, ranking, score-based ranking, learning-to-rank, responsible data management, responsible AI

ACM Reference Format:
Julia Stoyanovich, Meike Zehlike, and Ke Yang. 2023. Fairness in Ranking: From Values to Technical Choices and Back. In *Companion of the 2023 International Conference on Management of Data (SIGMOD-Companion '23), June 18–23, 2023, Seattle, WA, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3555041.3589405

1 INTRODUCTION

In the past few years, there has been much work on incorporating fairness requirements into the design of algorithmic rankers. And while numerous surveys on fairness in machine learning have been published, they typically focus on classification [4, 12, 32, 35]. In this tutorial, we give an overview of the large and growing body of work on fairness in raking, based on a two-part survey that we published in ACM Computing Surveys in 2022 [51, 52].

We consider two types of ranking tasks — score-based and supervised learning — and discuss how fairness has been operationalized for both. In score-based ranking, a given set of candidates is sorted on the score attribute, which may itself be computed on the fly, and returned in sorted order. In supervised learning-to-rank, a preference-enriched training set of candidates is given, with preferences among them stated in the form of scores, preference pairs, or lists; this training set is used to train a model that predicts the ranking of unseen candidates. For both score-based and learning-to-rank, we typically return the best-ranked k candidates, the top-k. Set selection is a special case of ranking that ignores the relative order among the top-k, returning them as a set.

While supervised learning-to-rank appears to be similar to classification, there is one crucial difference. The goal of classification is to assign a class label to each item, and this assignment is made independently for each item. In contrast, learning-to-rank positions items relative to each other, and so the outcome for one item is not independent of the outcomes for the other items. This lack of independence has profound implications for the design of learning-to-rank methods and, in particular, for fair learning-to-rank.

To make our discussion of fairness in ranking concrete, we now present an example from university admissions, a domain in which ranking and set selection are very natural and are broadly used.

1.1 Running example

Consider a university admissions officer who selects candidates from a large applicant pool. When making their decision, the officer pursues some or all of the goals listed below. Some of these may be legally mandated, while others may be based on the policies adopted by the university, and include admitting students who:

- are likely to succeed: complete the program with high marks and graduate on time;
- show strong interest in specific majors like computer science, art, or literature; and
- form a demographically diverse group in terms of their demographics, both overall and in each major.

Figure 1 shows a dataset C of applicants and illustrates the admissions process. Each applicant submits several quantitative scores, all transformed here to a discrete scale of 1 (worst) through 5 (best) for ease of exposition: X_1 is the high school GPA (grade point average), X_2 is the verbal portion of the SAT (Scholastic Assessment Test) score, and X_3 is the mathematics portion of the SAT score. Attribute X_4 (choice) is a weighted feature vector extracted from the applicant's essay, with weight ranging between 0 and 1, and with a higher value corresponding to stronger interest in a specific major. For example, candidate b is a White male with a high GPA (4 out of 5), perfect SAT verbal and SAT math scores (5 out of 5), a strong interest in studying computer science (feature weight 0.9), and some interest in studying art (weight 0.2).

The admissions officer uses a suite of tools to sift through the applications and identify promising candidates. These tools include

SIGMOD-Companion '23, June 18–23, 2023, Seattle, WA, USA

ACM ISBN 978-1-4503-9507-6/23/06.
https://doi.org/10.1145/3555041.3589405

candidate	A_1	A_2	X_1	X_2	X_3	X_4	Y_1	Y_2	Y_3
b	male	White	4	5	5	{cs:0.9; art:0.2}	14	9	1
c	male	Asian	5	3	4	{math:0.9; cs:0.5}	12	9	1
d	female	White	5	4	2	{lit:0.8; math:0.8}	11	4	6
e	male	White	3	3	4	{math:0.8; econ:0.4}	10	7	6
f	female	Asian	3	2	3	{econ:0.9; math:0.5}	8	5	8
k	female	Black	2	2	3	{lit:0.9;art:0.8}	7	1	9
l	male	Black	1	1	4	{lit:0.5; math:0.7}	6	6	2
o	female	White	1	1	2	{econ:0.9; cs:0.8}	4	7	8

(a)

τ_1
b
c
d
e
f
k
l
o

(b)

τ_2
c
b
e
f
d
o
l
k

(c)

τ_3
k
l
b
d
e
f
c
o

(d)

Figure 1: (a) dataset C of college applicants, with demographic attributes A_1 (sex) and A_2 (race), numerical attributes X_1 (high school GPA), X_2 (verbal SAT), and X_3 (math SAT), and attribute X_4 (choice) that is a vector extracted from the applicants' essays; (b) is a ranking τ_1 on Y_1, computed as the sum of X_1, X_2, and X_3; (c) is a ranking on Y_2, predicted based on historical performance of STEM (`cs`, `econ`, `math`) majors; (d) is a ranking on Y_3, predicted based on historical performance of humanities (`art`, `lit`) majors. In all cases, the top-4 candidates will be interviewed in score order, and potentially admitted.

score-based rankers that compute the score of each candidate based on a formula that the admissions officer gives, and then return some number of highest-scoring applicants in ranked order. This *scoring formula* may, for example, specify the score as a linear combination of the applicant's high school GPA and the two components of their SAT score, each carrying an equal weight. This is done in Figure 1(a), where a candidate's score is computed as $Y_1 = X_1 + X_2 + X_3$ and then ranking τ_1 in Figure 1(b) is produced.

Predictive analytics are also among the admissions officer's toolkits. For example, multiple ranking models may be trained (e.g., using any available learning-to-rank methods, such as RankNet [7] or ListNet [8]), one per undergraduate major or set of majors, on features X_1, X_2, X_3, X_4 of the successful applicants from the past years, to predict applicant's standing upon graduation (based, e.g., on their GPA in the major). These ranking models are then used to predict a ranking of this year's applicants. In our example in Figure 1(a), feature Y_2 predicts performance in a STEM major such as computer science (`cs`), economics (`econ`), or mathematics (`math`) and leads to ranking τ_2 in Figure 1(c), while feature Y_3 predicts performance in a humanities major such as literature (`lit`) or fine arts (`art`) and leads to ranking τ_3 in Figure 1(d). The promising applicants identified in this way—with the help of either a score-based ranker or a predictive analytic—will then be considered more closely, *in ranked order*: invited for an interview and potentially admitted.

Let us recall that, in addition to incorporating quantitative scores and students' choices, an admissions officer also aims to admit a demographically diverse group of students to the university and to each major. Further, the admissions officer is increasingly aware that the data on which their decisions are based may be biased, in the sense that this data may carry results of historical discrimination or disadvantage [38], and that the computational tools at their disposal may be exacerbating or introducing new forms of bias, or even creating a kind of a self-fulfilling prophecy. For this reason, the officer may elect to incorporate one or several fairness objectives into the ranking process.

For example, they may assert, for legal or ethical reasons, that the proportion of the female applicants among those selected for further consideration should match their proportion in the input. Further, the admissions officer may assert that, because applicants are interviewed in ranked order, it is important to achieve proportional representation by sex in *every prefix* of the produced ranking. In this tutorial, we give an overview of the technical work that would allow an admissions officer to compute ranked results under these and other fairness requirements.

1.2 Scope and contributions

We are aware of several recent tutorials on fairness in ranking at SIGIR 2019 [9], RecSys 2019 [17], VLDB 2020 [2], and ICDE 2021 [36], covering different approaches and pointing to the need to systematize the work on fairness in ranking. In this tutorial, we offer a broad perspective, connecting work across subfields. In the remainder of this document, we give an overview of the content of the tutorial, and refer the reader to the survey on which this tutorial is based for additional details [51, 52].

2 CLASSIFICATION FRAMEWORK: RECONCILING VALUES WITH TECHNICAL CHOICES

Which specific fairness requirements a decision maker will assert depends on the values they are operationalizing and, thus, on the mitigation objectives. An important goal of our tutorial is to present a classification framework for fair ranking methods that helps establish the correspondence between normative dimensions and technical design choices. Figure 2 presents this classification framework as a mind map.

Operationally, algorithmic approaches to fair ranking differ in how they represent candidates (e.g., whether they support one or multiple sensitive attributes, and whether these are binary), in what fairness measure(s) they adopt, in how they navigate the trade-offs between fairness and utility during mitigation, and at what processing stage a mitigation is applied. Conceptually, these operational choices correspond to normative statements about the types of bias

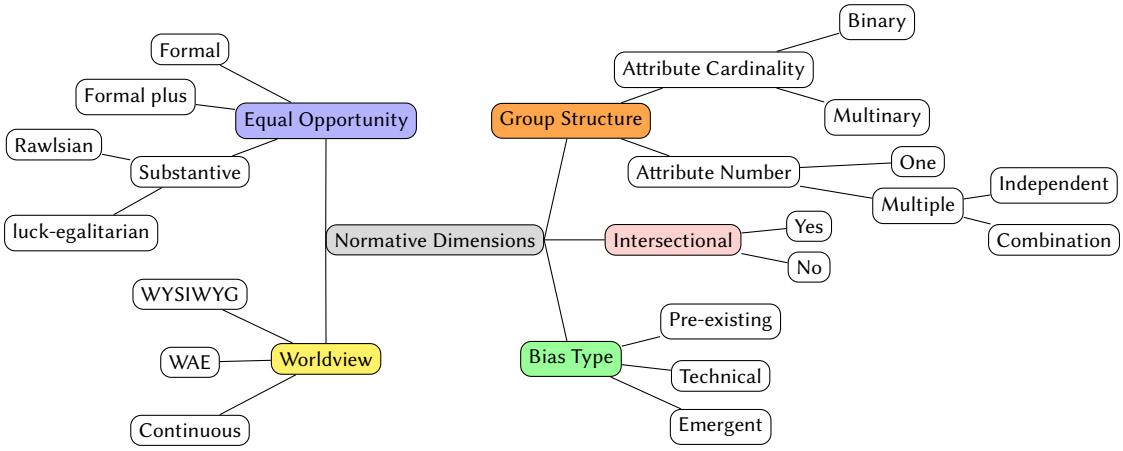

Figure 2: A mind map summary of the classification framework for fair ranking methods.

Table 1: Summary of fair score-based ranking methods.

Method	Group structure	Bias	Worldview	EOP	Intersectional
Rank-aware proportional representation [46]	one binary sensitive attr.	pre-existing	WAE	luck-egalitarian	no
Constrained ranking maximization [11]	multiple sensitive attrs.; multinary; handled independently	pre-existing	WAE	luck-egalitarian (1 sensitive attr. only)	no
Balanced diverse ranking [44]	multiple sensitive attrs.; multinary; handled independently	pre-existing; technical	WAE	luck-egalitarian	yes
Diverse k-choice secretary [43]	one multinary sensitive attr.	pre-existing	WAE	luck-egalitarian	no
Utility of selection with implicit bias [28]	one binary sensitive attr.	pre-existing; implicit	WAE	N/A	no
Utility of ranking with implicit bias [10]	multiple sensitive attrs.; multinary; handled independently	pre-existing; implicit	WAE	N/A	yes
Causal intersectionally fair ranking [45]	multiple sensitive attrs.; multinary; handled independently	pre-existing	WAE	Rawlsian	yes
Designing fair ranking functions [1]	any	pre-existing	any	any	yes

being observed and mitigated, and to the mitigation objectives. We now summarize the facets of the classificaiton framework.

Group structure. Fairness of a method is commonly stated with respect to a set of categorical sensitive attributes (or features). We discuss several orthogonal dimensions of group structure, based on the handling of sensitive attributes. Some methods consider only *binary* sensitive attributes, while other methods handle higher-cardinality domains of sensitive attribute values. If a higher-cardinality domain is supported, methods differ in whether they consider one of the values to be protected (a single designated group that has been experiencing discrimination), or if they treat all sensitive attribute values as potentially being subject to discrimination.

Further, some methods are designed to handle a *single sensitive attribute* at a time (e.g., they handle either gender or race), while other methods handle *multiple sensitive attributes* simultaneously (e.g., they handle both gender and race). Methods that support multiple sensitive attributes differ in whether they handle these *independently* (e.g., by asserting fairness constraints w.r.t. the treatment of both women and Blacks) or *in combination* (e.g., by requiring fairness w.r.t. Black women).

Intersectional discrimination. Intersectional Discrimination [14, 30] states that individuals who belong to several protected groups simultaneously (e.g., Black women) experience stronger discrimination compared to individuals who belong to a single protected

Table 2: Summary of fair learning-to-rank methods.

Method	Mitigation Point	Group structure	Bias	Worldview	EOP Framework
iFair [29]	pre-proc.	multiple multinary attr.; independent	technical	WYSWYG	formal
DELTR [48]	in-proc.	one binary attr.	pre-existing	WAE	luck-egalitarian
Fair-PG-Rank [42]	in-proc	one binary attr.	technical	WYSIWYG	formal
Pairwise Ranking Fairness [5]	in-proc.	one binary attr.	?	WYSIWYG	formal-plus
FA*IR [47] & [50]	post-proc.	one multinary attr.; combination	pre-existing	continuous	formal / luck-egalitarian
Fair Ranking at LinkedIn [22]	post-proc.	one multinary attr.; combination	pre-existing; technic	continuous	none / luck-egalitarian (1 sensitive attr.)
CFAθ [49]	post-proc.	multiple binary attr.; combination	pre-existing	continuous	formal / substantive
Fairness of Exposure [41]	post-proc.	one binary attr.	pre-existing/ technical	WYSIWYG / WAE	formal / luck-egalitarian
Equity of Attention [6]	post-proc.	one multinary attr.; independent	technical / emergent	WYSIWYG	formal

group (e.g., White women or Black men), and that this disadvantage compounds more than additively. This effect has been demonstrated by numerous case studies, and by theoretical and empirical work [13, 15, 33, 40]. An immediate interpretation for ranking is that, if fairness is taken to mean proportional representation among the top-k, then it is possible to achieve proportionality for each gender subgroup (e.g., men and women) and for each racial subgroup (e.g., Black, and White), while still having inadequate representation for a subgroup defined by the intersection of both attributes (e.g., Black women).

Intersectional concerns also arise in more subtle ways. For example, when constraints are stated on individual attributes, like race and gender, and the goal is to maximize score-based utility subject to these constraints, then a particular kind of unfairness can arise: utility loss can be particularly severe in historically disadvantaged intersectional groups [44].

Type of bias. that a fair ranking method attempts to mitigate — pre-existing, technical bias [3, 6] or emergent bias [34], as defined by [20] — is another important technical dimension with far-reaching normative consequences. We give examples of how each type of bias may arise in ranking, and classify fair ranking methods based on which bias type they aim to mitigate.

Mitigation objectives. This is a rich normative dimension of our classification framework that includes both the *worldviews* framing of Friedler et al. [19], and the recently-proposed re-interpretation of *equality of opportunity (EO) doctrines* for algorithmic fairness by Arif Khan et al. [27].

We classify some fair ranking methods as those that are consistent with formal EO, interpreted as either fairness through blindness or formal-plus EO [18]. These methods require calibrated performance across groups [24, 28]. We classify other fair ranking methods as those that are consistent with substantive EO. These are, in turn, subdivided into backward-facing (i.e., correcting for a history of disadvantage and taking the luck-egalitarian perspective [16, 39]) and forward-facing (i.e., ensuring equitable access to opportunity over a lifetime and taking the Rawls' Fair EO perspective [37]).

3 FAIRNESS IN SCORE-BASED RANKING

In score-based ranking, we categorize mitigation methods into those that intervene on the score distribution, on the scoring function, or on the ranked outcome. Methods that *intervene on the score distribution* aim to mitigate disparities in candidate scores, either before these candidates are processed by an algorithmic ranker or during ranking. Methods that *intervene on the ranking function* identify a function that is similar to the input function but that produces a ranked outcome that meets the specified fairness criteria. Methods that *intervene on the ranked outcome* impose constraints to require a specific level of diversity or representation among the top-k as a set, or in every prefix of the top-k.

We present a selection of approaches for fairness in score-based ranking listed in Table 1. All methods we present are mapped to our classification framework, bringing out their commonalities and differences that go beyond the purely technical choices, and allowing us to reason about trade-offs.

4 FAIRNESS IN LEARNING-TO-RANK

In supervised learning, we categorize mitigation methods into pre-processing, in-processing, and post-processing. *Pre-processing* methods seek to mitigate discriminatory bias in training data, and have the advantage of early intervention on the pre-existing bias. *In-processing* methods aim to learn a bias-free model. Finally, *post-processing* methods re-rank candidates in the output subject to given fairness constraints [23]. To mitigate unfairness, two main lines of work on fairness-enhancing interventions have also emerged over the past several years: probability-based [46, 47] and exposure-based [3, 26]. During the tutorial, we give an overview of the methods listed in Table 2.

5 PRACTICAL GUIDANCE

The final part of the tutorial consists of a discussion regarding the practical aspects of fair ranker design, and recommendations for the evaluation of fair ranking methods [25, 31].

For example, those methods that explicitly incorporate a notion of utility *into their fairness objective*, namely Biega et al. [6], Lahoti et al. [29], Singh and Joachims [42], and the disparate treatment and disparate impact definition of Singh and Joachims [41]) generally lean towards the WYSIWYG worldview and are consistent with formal EO. In contrast, methods that explicitly exclude a utility measure from the fairness definition (Geyik et al. [22], Zehlike et al. [47], Zehlike and Castillo [48], Zehlike et al. [49], and the demographic parity definition of Singh and Joachims [41]), generally lean towards the WAE worldview and substantive EO. Additionally, some methods explicitly allow continuous interpolation between two worldviews WAE and WYSIWYG, either by introducing a sliding parameter or by allowing a range of values for the fairness constraints (Geyik et al. [22], Zehlike et al. [47], Zehlike and Castillo [48], Zehlike et al. [49]). With these recommendations, we aim to establish best practices for the development, evaluation, and deployment of fair ranking algorithms, and to avoid potentially harmful uninformed transfer of methods between application domains.

As an interactive component of the final portion of the tutorial, we discuss the fair ranking method by García-Soriano and Bonchi [21]. This method proposes to trade off the WAE and WYSIWYG worldviews in a specific way. The goal of the discussion is to situate this method within the classification framework of Section 2, and to compare it with some of the other surveyed methods based on the normative dimensions that are induced by the technical choices.

6 PRESENTER BIOS

Prof. Julia Stoyanovich is Institute Associate Professor of Computer Science & Engineering, Associate Professor of Data Science, and Director of the Center for Responsible AI at New York University. Julia's goal is to make "Responsible AI" synonymous with "AI." She works towards this goal by engaging in academic research, education, and technology policy, and by speaking about the benefits and harms of AI to practitioners and members of the public. Julia's research interests include AI ethics and legal compliance, data management and AI systems, and computational social choice. She developed and teaches technical courses on responsible data science, co-developed courses We are AI: Taking Control of Technology and AI Ethics: Global Perspectives, and co-created two comic book series on responsible AI. Julia is engaged in technology policy in the US and internationally, having served on the New York City Automated Decision Systems Task Force, by mayoral appointment, among other roles. She received M.S. and Ph.D. degrees in Computer Science from Columbia University, and a B.S. degree in Computer Science and in Mathematics & Statistics from the University of Massachusetts at Amherst. Julia is a recipient of an NSF CAREER award and is a Senior Member of the ACM.

Dr. Meike Zehlike is a Senior Applied Scientist in the Algorithmic Privacy and Fairness team at Zalando Research, and an ethical AI consultant. She is currently visiting the Complexity Science Hub in Vienna as a Postdoctoral Researcher. She earned her Ph.D. in computer science at Humboldt Universität zu Berlin in 2022, working under Ulf Leser (HU), Carlos Castillo (UPF Barcelona), and Krishna Gummadi (MPI-SWS Saarbrücken) on a dissertation on "Fair rankings." Meike received several grants and awards for her Ph.D. research, including the Data Transparency Lab Research grant and the Humboldt Universität Best Dissertation Award. Her interests center around artificial intelligence and its social impacts, translating ethics into math, and operationalizing ethical AI principles in e-commerce contexts.

Dr. Ke Yang is a Postdoctoral Researcher at the University of Massachusetts at Amherst. Her research centers around data management and machine learning, including algorithmic fairness, model explanation, and the social impacts of algorithms in data-driven systems. Ke received her Ph.D. from the Tandon School of Engineering at New York University in 2021, where she was advised by Julia Stoyanovich. She received the Pearl Brownstein Doctoral Research Award, given to Computer Science & Engineering Ph.D. students at NYU Tandon whose doctoral research shows the greatest promise, for her dissertation on "Fairness, diversity, and interpretability in ranking." Ke received her B.E. and M.S. degrees in 2012 and 2015, respectively, from Beijing Technology and Business University in China.

7 ACKNOWLEDGEMENTS

This research was supported in part by NSF awards No. 1934464, 1916505, and 1922658.

REFERENCES

[1] Abolfazl Asudeh, HV Jagadish, Julia Stoyanovich, and Gautam Das. 2019. Designing fair ranking schemes. In *Proceedings of the 2019 International Conference on Management of Data*. ACM, 1259–1276.

[2] Abolfazl Asudeh and H. V. Jagadish. 2020. Fairly Evaluating and Scoring Items in a Data Set. *Proc. VLDB Endow.* 13, 12 (2020), 3445–3448. https://doi.org/10.14778/3415478.3415566

[3] Ricardo Baeza-Yates. 2018. Bias on the web. *Commun. ACM* 61, 6 (2018), 54–61. https://doi.org/10.1145/3209581

[4] Solon Barocas, Moritz Hardt, and Arvind Narayanan. 2017. Fairness in machine learning. *Nips tutorial* 1 (2017), 2.

[5] Alex Beutel, Jilin Chen, Tulsee Doshi, Hai Qian, Li Wei, Yi Wu, Lukasz Heldt, Zhe Zhao, Lichan Hong, Ed H. Chi, and Cristos Goodrow. 2019. Fairness in Recommendation Ranking Through Pairwise Comparisons. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (Anchorage, AK, USA) *(KDD '19)*. ACM, New York, NY, USA, 2212–2220. https://doi.org/10.1145/3292500.3330745

[6] Asia J Biega, Krishna P Gummadi, and Gerhard Weikum. 2018. Equity of attention: Amortizing individual fairness in rankings. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. ACM, 405–414.

[7] Christopher JC Burges. 2010. From ranknet to lambdarank to lambdamart: An overview. *Learning* 11, 23-581 (2010), 81.

[8] Zhe Cao, Tao Qin, Tie-Yan Liu, Ming-Feng Tsai, and Hang Li. 2007. Learning to rank: from pairwise approach to listwise approach. In *Proceedings of the 24th international conference on Machine learning*. ACM, 129–136.

[9] Carlos Castillo. 2019. Fairness and transparency in ranking. In *ACM SIGIR Forum*, Vol. 52. ACM New York, NY, USA, 64–71.

[10] L Elisa Celis, Anay Mehrotra, and Nisheeth K Vishnoi. 2020. Interventions for ranking in the presence of implicit bias. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. 369–380.

[11] L Elisa Celis, Damian Straszak, and Nisheeth K Vishnoi. 2018. Ranking with Fairness Constraints. In *45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

[12] Alexandra Chouldechova and Aaron Roth. 2020. A snapshot of the frontiers of fairness in machine learning. *Commun. ACM* 63, 5 (2020), 82–89. https://doi.org/10.1145/3376898

[13] Patricia Hill Collins. 2002. *Black feminist thought: Knowledge, consciousness, and the politics of empowerment.* routledge.

[14] Kimberle Crenshaw. 1990. Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stan. L. Rev.* 43 (1990), 1241.

[15] Catherine D'Ignazio and Lauren F Klein. 2020. *Data feminism.* MIT Press.

[16] Ronald Dworkin. 1981. What is Equality? Part 1: Equality of Welfare. *Philosophy and Public Affairs* 10, 3 (1981), 185–246. http://www.jstor.org/stable/2264894

[17] Michael D. Ekstrand, Robin Burke, and Fernando Diaz. 2019. Fairness and discrimination in recommendation and retrieval. 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, 576–577. https://doi.org/10.1145/3298689.3346964

[18] Joseph Fishkin. 2014. *Bottlenecks: A New Theory of Equal Opportunity.* Oup Usa.

[19] Sorelle A Friedler, Carlos Scheidegger, and Suresh Venkatasubramanian. 2016. On the (im) possibility of fairness. *arXiv preprint arXiv:1609.07236* (2016).

[20] Batya Friedman and Helen Nissenbaum. 1996. Bias in Computer Systems. *ACM Trans. Inf. Syst.* 14, 3 (1996), 330–347. https://doi.org/10.1145/230538.230561

[21] David García-Soriano and Francesco Bonchi. 2021. Maxmin-Fair Ranking: Individual Fairness under Group-Fairness Constraints. In *KDD '21: The 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event, Singapore, August 14-18, 2021*, Feida Zhu, Beng Chin Ooi, and Chunyan Miao (Eds.). ACM, 436–446. https://doi.org/10.1145/3447548.3467349

[22] Sahin Cem Geyik, Stuart Ambler, and Krishnaram Kenthapadi. 2019. Fairness-aware ranking in search & recommendation systems with application to linkedin talent search. In *Proceedings of the 25th acm sigkdd international conference on knowledge discovery & data mining*. 2221–2231.

[23] Sara Hajian, Francesco Bonchi, and Carlos Castillo. 2016. Algorithmic bias: From discrimination discovery to fairness-aware data mining. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. ACM, 2125–2126.

[24] Moritz Hardt, Eric Price, and Nati Srebro. 2016. Equality of opportunity in supervised learning. In *Advances in neural information processing systems*. 3315–3323.

[25] Kalervo Järvelin and Jaana Kekäläinen. 2002. Cumulated gain-based evaluation of IR techniques. *ACM Transactions on Information Systems (TOIS)* 20, 4 (2002), 422–446.

[26] Thorsten Joachims, Laura Granka, Bing Pan, Helene Hembrooke, and Geri Gay. 2017. Accurately interpreting clickthrough data as implicit feedback. In *ACM SIGIR Forum*, Vol. 51. Acm New York, NY, USA, 4–11.

[27] Falaah Arif Khan, Eleni Manis, and Julia Stoyanovich. 2022. Towards Substantive Conceptions of Algorithmic Fairness: Normative Guidance from Equal Opportunity Doctrines. In *Equity and Access in Algorithms, Mechanisms, and Optimization, EAAMO 2022, Arlington, VA, USA, October 6-9, 2022*. ACM, 18:1–18:10. https://doi.org/10.1145/3551624.3555303

[28] Jon Kleinberg and Manish Raghavan. 2018. Selection Problems in the Presence of Implicit Bias. In *9th Innovations in Theoretical Computer Science Conference (ITCS 2018) (Leibniz International Proceedings in Informatics (LIPIcs), Vol. 94)*, Anna R. Karlin (Ed.). Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 33:1–33:17. https://doi.org/10.4230/LIPIcs.ITCS.2018.33

[29] Preethi Lahoti, Krishna P Gummadi, and Gerhard Weikum. 2019. ifair: Learning individually fair data representations for algorithmic decision making. In *2019 IEEE 35th International Conference on Data Engineering (ICDE)*. IEEE, 1334–1345.

[30] Timo Makkonen. 2002. Multiple, compound and intersectional discrimination: Bringing the experiences of the most marginalized to the fore. *Institute for Human Rights, Åbo Akademi University* (2002).

[31] Christopher D Manning, Prabhakar Raghavan, and Hinrich Schütze. 2008. Evaluation in information retrieval. *Introduction to information retrieval* 1 (2008), 188–210.

[32] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. 2021. A Survey on Bias and Fairness in Machine Learning. *ACM Comput. Surv.* 54, 6 (2021), 115:1–115:35. https://doi.org/10.1145/3457607

[33] Safiya Umoja Noble. 2018. *Algorithms of oppression: How search engines reinforce racism.* nyu Press.

[34] Bing Pan, Helene Hembrooke, Thorsten Joachims, Lori Lorigo, Geri Gay, and Laura Granka. 2007. In Google we trust: Users' decisions on rank, position, and relevance. *Journal of computer-mediated communication* 12, 3 (2007), 801–823.

[35] Dana Pessach and Erez Shmueli. 2020. Algorithmic Fairness. *CoRR* abs/2001.09784 (2020). arXiv:2001.09784 https://arxiv.org/abs/2001.09784

[36] Evaggelia Pitoura, Kostas Stefanidis, and Georgia Koutrika. 2021. Fairness in Rankings and Recommenders: Models, Methods and Research Directions. In *37th IEEE International Conference on Data Engineering, ICDE 2021, Chania, Greece, April 19-22, 2021*. IEEE, 2358–2361. https://doi.org/10.1109/ICDE51399.2021.00265

[37] John Rawls. 1971. *A theory of justice.* Harvard University Press.

[38] Richard V. Reeves and Dimitrios Halikias. 2017. Race gaps in SAT scores highlight inequality and hinder upward mobility. (2017). https://www.brookings.edu/research/race-gaps-in-sat-scores-highlight-inequality-and-hinder-upward-mobility

[39] John E. Roemer. 2002. Equality of opportunity: a progress report. *Social Choice and Welfare* 19, 2 (2002), 405–471.

[40] Stephanie A Shields. 2008. Gender: An intersectionality perspective. *Sex roles* 59, 5-6 (2008), 301–311.

[41] Ashudeep Singh and Thorsten Joachims. 2018. Fairness of exposure in rankings. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, 2219–2228.

[42] Ashudeep Singh and Thorsten Joachims. 2019. Policy Learning for Fairness in Ranking. *arXiv preprint arXiv:1902.04056* (2019).

[43] Julia Stoyanovich, Ke Yang, and H. V. Jagadish. 2018. Online Set Selection with Fairness and Diversity Constraints. In *Proceedings of the 21th International Conference on Extending Database Technology, EDBT 2018, Vienna, Austria, March 26-29, 2018*. 241–252. https://doi.org/10.5441/002/edbt.2018.22

[44] Ke Yang, Vasilis Gkatzelis, and Julia Stoyanovich. 2019. Balanced Ranking with Diversity Constraints. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019*. 6035–6042. https://doi.org/10.24963/ijcai.2019/836

[45] Ke Yang, Joshua R. Loftus, and Julia Stoyanovich. 2021. Causal intersectionality and fair ranking. In *Symposium on Foundations of Responsible Computing (FORC)*. https://doi.org/10.4230/LIPIcs.FORC.2021.7

[46] Ke Yang and Julia Stoyanovich. 2017. Measuring Fairness in Ranked Outputs. In *Proceedings of the 29th International Conference on Scientific and Statistical Database Management*. ACM, 22.

[47] 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*. ACM, 1569–1578.

[48] Meike Zehlike and Carlos Castillo. 2018. Reducing disparate exposure in ranking: A learning to rank approach. *arXiv preprint arXiv:1805.08716* (2018).

[49] Meike Zehlike, Philipp Hacker, and Emil Wiedemann. 2017. Matching code and law: achieving algorithmic fairness with optimal transport. *Data Mining and Knowledge Discovery* (2017), 1–38.

[50] Meike Zehlike, Tom Sühr, Ricardo Baeza-Yates, Francesco Bonchi, Carlos Castillo, and Sara Hajian. 2022. Fair Top-k Ranking with multiple protected groups. *Information Processing & Management* 59, 1 (2022), 102707.

[51] Meike Zehlike, Ke Yang, and Julia Stoyanovich. 2022. Fairness in Ranking, Part I: Score-Based Ranking. *ACM Comput. Surv.* (apr 2022). https://doi.org/10.1145/3533379

[52] Meike Zehlike, Ke Yang, and Julia Stoyanovich. 2022. Fairness in Ranking, Part II: Learning-to-Rank and Recommender Systems. *ACM Comput. Surv.* (apr 2022). https://doi.org/10.1145/3533380