# The Winner Takes it All: Geographic Imbalance and Provider (Un)fairness in Educational Recommender Systems

Elizabeth Gómez
Universitat de Barcelona
Barcelona, Spain
egomezye13@alumnes.ub.edu

Carlos Shui Zhang
Universitat de Barcelona
Barcelona, Spain
cshuizha7@alumnes.ub.edu

Ludovico Boratto*
University of Cagliari
Cagliari, Italy
ludovico.boratto@acm.org

Maria Salamó
Universitat de Barcelona
Barcelona, Spain
maria.salamo@ub.edu

Mirko Marras
EPFL
Lausanne, Switzerland
mirko.marras@acm.org

## ABSTRACT

Educational recommender systems channel most of the research efforts on the effectiveness of the recommended items. While teachers have a central role in online platforms, the impact of recommender systems for teachers in terms of the exposure such systems give to the courses is an under-explored area. In this paper, we consider data coming from a real-world platform and analyze the distribution of the recommendations w.r.t. the geographical provenience of the teachers. We observe that data is highly imbalanced towards the United States, in terms of offered courses and of interactions. These imbalances are exacerbated by recommender systems, which overexpose the country w.r.t. its representation in the data, thus generating unfairness for teachers outside that country. To introduce equity, we propose an approach that regulates the share of recommendations given to the items produced in a country (*visibility*) and the position of the items in the recommended list (*exposure*).

## CCS CONCEPTS

• **Information systems → Learning to rank**; **Recommender systems**; **Personalization**; **Collaborative filtering**; • **Applied computing → Law, social and behavioral sciences**.

## KEYWORDS

MOOC, Education, Online Course, Fairness, Bias, Data Imbalance.

**ACM Reference Format:**
Elizabeth Gómez, Carlos Shui Zhang, Ludovico Boratto, Maria Salamó, and Mirko Marras. 2021. The Winner Takes it All: Geographic Imbalance and Provider (Un)fairness in Educational Recommender Systems. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR ’21), July 11–15, 2021, Virtual Event, Canada.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3404835.3463235

*Work partially done while at EURECAT - Centre Tecnològic de Catalunya.


*SIGIR ’21, July 11–15, 2021, Virtual Event, Canada*

ACM ISBN 978-1-4503-8037-9/21/07…$15.00
https://doi.org/10.1145/3404835.3463235

## 1 INTRODUCTION

Learning paradigms are shifting towards online environments [15, 27], thanks to Massive Online Open Courses (MOOC) platforms. Recommender systems are the means that allows MOOC platforms to direct appropriate resources to learners [4]. Course recommendation functionalities are common in these platforms, with a clear focus on the learners and the opportunities offered to them [19].

The impact that recommender systems have on teachers is an under-explored perspective. However, teachers are a key stakeholders in a MOOC platform, since they are the ones that provide the courses, and they are directly affected by the way recommendation lists are shaped. Indeed, according to how many times the courses of a teacher are recommended (*visibility*) [10] and where they appear in the ranking, that teacher is given a certain *exposure* by the system [28]. Disparities in the visibility and exposure given to teachers might lead to undesired consequences [24], such as unfairness [11, 23]. In this paper, we focus on *group unfairness*, shaping groups based on the *geographic provenience* of the teachers offering the courses. Our goal is to study if imbalances in the country of provenience of the teachers might affect the opportunities of teachers coming from certain parts of the world to offer their services, by being under-exposed. Specifically, we consider two demographic groups, the first covering the country with the highest representation of teachers in the platform (in our data, the United States), and the second containing the rest of the world. There are multiple reasons why this is an interesting setting. Considering the reference dataset for this study, COCO [9] (presented in Section 2.4), the United States covers more than 40% of the courses and nearly 50% of the ratings. The remaining 73 countries attract a very small percentage of ratings and courses, thus leading to an important *geographic imbalance* in the input data. However, in a binary setting such as the one we consider, the most represented country does not constitute an overall majority in the data. This offers an interesting benchmark to study the interplay between geographic imbalance and minority groups and their impact on unfairness.

When recommender systems overexpose teachers coming from the country with the highest representation, teachers from the rest of the world are unfairly affected by how recommendations are generated. In this work, we consider five state-of-the-art collaborative filtering models, and show that they exacerbate disparities emerging from geographic imbalance, under-exposing the teachers coming from the rest of the world. To overcome these phenomena, we propose a re-ranking approach that aims to re-distribute the recommendations between the United States and the rest of the world, following a notion of *equity* [30].

Specifically, our contributions are as follows: (*i*) we assess how recommender systems affect groups of teachers based on their provenience, (*ii*) we propose an approach to introduce equity in the recommendations' distribution, and (*iii*) we show that we can introduce equity without affecting recommendation effectiveness.

# 2 PRELIMINARIES

## 2.1 Recommendation scenario

Let $U = \{u_1, u_2, ..., u_n\}$ be a set of learners, $C = \{c_1, c_2, ..., c_j\}$ be a set of courses, and $V$ be a totally ordered set of values used to express a preference. The set of ratings is a ternary relation $R \subseteq U \times C \times V$; each rating is denoted by $r_{uc}$. We consider a temporal split of the data, where a fixed percentage of the ratings of the learners (ordered by timestamp) is used for training and the rest goes to the test set [1].

The recommendation goal is to learn a function $f$ that estimates the relevance ($\hat{r}_{uc}$) of the learner-course pairs that do not appear in the training data. We denote as $\hat{R}$ the set of recommendations, and as $\hat{R}_G$ those involving courses of a group $G$.

Let $A = \{a_1, a_2, ..., a_k\}$ be the set of geographic areas in which courses are organized. Specifically, we consider a geographic area as the country associated to a course. We denote as $A_c$ the set of geographic areas of a course $c$. Note that, since teachers of a course could be from different geographical areas, several geographic areas may appear in a course. We shape two groups, the most represented area, $M = \{c \in C : 1 \in A_c\}$, and the rest, $m = \{c \in C : 1 \notin A_c\}$. Note that 1 identifies the most represented country.

## 2.2 Metrics

**Representation.** The representation of a group is the amount of times that this group appears in the data. We consider two forms of representation, based on (*i*) the amount of courses offered by a group and (*ii*) the amount of ratings collected for that group. We define with $\mathcal{R}$ the *representation* of a group $G$ ($\mathcal{R}_C$ denotes a course-based representation, while $\mathcal{R}_R$ a rating-based representation):

$$\mathcal{R}_C(G) = |G|/|C| \quad (1)$$

$$\mathcal{R}_R(G) = |\{r_{uc} : c \in G\}|/|R| \quad (2)$$

Eq. (1) accounts for the proportion of courses of a group, while Eq. (2) for the proportion of ratings associated to a group. The representation of a group is measured by considering only the training set. Given a group $G$, the representation of the other, $\overline{G}$, is computed as $\mathcal{R}_*(\overline{G}) = 1 - \mathcal{R}_*(G)$ (where '*' refers to $C$ or $R$).

**Disparate Impact.** We assess unfairness with two notions of *disparate impact* generated by a recommender system.

*Definition 2.1 (Disparate visibility).* The *disparate visibility* of a group is the difference between the share of recommendations for items of that group and the representation of that group [10]:

$$\Delta\mathcal{V}(G) = \frac{1}{|U|}\sum_{u \in U}\frac{|\{\hat{r}_{uc} : c \in \hat{R}_G\}|}{|\hat{R}|} - \mathcal{R}_*(G) \quad (3)$$

Its range is in $[-\mathcal{R}_*(G), 1-\mathcal{R}_*(G)]$; it is 0 when there is no disparate visibility, while negative/positive values indicate that the group had a share of recommendations lower/higher than its representation.

*Definition 2.2 (Disparate exposure).* The *disparate exposure* of a group is the difference between the exposure obtained by the group in the recommendations [28] and the representation of that group:

$$\Delta\mathcal{E}(G) = \frac{1}{|U|}\sum_{u \in U}\frac{\sum_{pos=1}^{k}\frac{1}{log_2(pos+1)}, \forall c \in \hat{R}_G}{\sum_{pos=1}^{k}\frac{1}{log_2(pos+1)}} - \mathcal{R}_*(G) \quad (4)$$

where *pos* is the position of an item in the top-$k$ recommendations.

This metric also ranges in $[-\mathcal{R}_*(G), 1-\mathcal{R}_*(G)]$; it is 0 when there is no disparate exposure, while negative/positive values indicate that the exposure given to the group in the recommendations is lower/higher than its representation.

Notice that the disparate visibility/exposure of one group can be computed as the opposite of the value obtained for the other group.

## 2.3 Recommendation algorithms

We consider five Collaborative Filtering algorithms. For the class of memory-based models, we choose UserKNN [12] and ItemKNN [26]. As matrix factorization based approaches, we consider BPR [25], BiasedMF [17], and SVD++ [16]. We contextualize our results with two non-personalized models (Most Popular and Random Guess).

## 2.4 Dataset

To the best of our knowledge, COCO [9] is the only educational dataset that contains the geographic provenience of the users. It was collected from an online course platform, and each course is associated to one or more teachers, belonging to 74 countries.

We pre-processed the dataset to remove all learners with less than 3 ratings. Our final dataset contains 12,472 courses and 298,644 learners, who provided 1,296,598 ratings. We encoded each country with subsequent integers, with the United States having ID 1.

# 3 DISPARATE IMPACT ASSESSMENT

## 3.1 Experimental setting

The test is composed by the most recent 20% of the ratings of each learner. We run the recommendation algorithms using the LibRec library (v.2). For each user, we store the first 100 results (top-$n$) to then mitigate disparities through a re-ranking. The recommendation list for each learner is composed by 20 courses (top-$k$).

Each algorithm was run with the following hyper-parameters: (*i*) **UserKNN.** similarity: Pearson; neighbors: 50; similarity shrinkage: 10; (*ii*) **ItemKNN.** similarity: Cosine; neighbors: 200; similarity shrinkage: 10; (*iii*) **BPR.** iterator learnrate: 0.01; iterator learnrate maximum: 0.01; iterator maximum: 100; user regularization: 0.01; item regularization: 0.01; factor number: 10; learnrate bolddriver: false; learnrate decay=1.0; (*iv*) **BiasedMF.** iterator learnrate: 0.01; iterator learnrate maximum: 0.01; iterator maximum: 10; user regularization: 0.01; item regularization: 0.01; bias regularization: 0.01; number of factors: 10; learnrate bolddriver: false; learnrate decay: 1.0; (*v*) **SVD++.** iterator learnrate: 0.01; iterator learnrate maximum: 0.01; iterator maximum: 13; user regularization: 0.01; item regularization: 0.01; impItem regularization: 0.001; number of factors: 10; learnrate bolddriver: false; learnrate decay: 1.0.

**Table 1: Effectiveness, disparate visibility, and disparate exposure of group $m$, considering both a course- and a rating-based representation of the groups.**

| Algorithm | NDCG | $\Delta\mathcal{V}_C$ | $\Delta\mathcal{E}_C$ | $\Delta\mathcal{V}_R$ | $\Delta\mathcal{E}_R$ |
|---|---|---|---|---|---|
| **MostPop** | 0.0193 | -0.3091 | -0.2117 | -0.2447 | -0.1473 |
| **RandomG** | 0.0006 | **0.0000** | **-0.0001** | 0.0644 | 0.0643 |
| **UserKNN** | 0.0372 | -0.0402 | -0.1457 | 0.0242 | -0.0813 |
| **ItemKNN** | **0.2068** | -0.0862 | -0.0783 | -0.0218 | -0.0139 |
| **BPR** | 0.1401 | -0.0715 | -0.0658 | **-0.0071** | **-0.0014** |
| **BiasedMF** | 0.0007 | -0.1065 | -0.0949 | -0.0421 | -0.0305 |
| **SVD++** | 0.0044 | -0.0534 | -0.0543 | 0.0110 | 0.0101 |

To evaluate recommendation quality, we measure the NDCG [13].

## 3.2 Characterizing User Behavior

In COCO, $\mathcal{R}_C(M) = 0.41$ and $\mathcal{R}_R(M) = 0.47$. Considering that the dataset contains 74 countries, we observe a strong geographic imbalance in terms of offered courses. This imbalance is worsened when we consider the ratings. The vast majority of the ratings in this dataset is 5 [4]. Also under this geographical setting, user satisfaction is equally distributed along the two groups.

**Observation 1**. *There is a strong geographic imbalance in the representation of each group, in terms of offered items. The most represented group usually attracts more ratings, thus increasing the existing imbalance.*

## 3.3 Assessing Effectiveness and Disparities

In Table 1, we report the results obtained by each model. Results show that ItemKNN is the most effective algorithm. Considering that the rating distribution is skewed towards high values, these results connect to widely-known phenomena that make the algorithm successful [21], such as the data size and the fact that the neighborhoods will not change much, given that the ratings are very similar. When considering a course-based representation, Random Guess provides the most equitable visibility and exposure. Hence, when picking the items to recommend at random, the recommendation lists are shaped following the distribution in the course offer; nevertheless, this is the least effective algorithm. Finally, BPR returns the most equitable recommendations when considering a rating-based representation. We connect these results to those of Cremonesi et al. [8], who showed that factorization approaches can recommend long-tail items, by building factors that capture all the preferences. BPR also returns the second best NDCG.

**Observation 2**. *Geographic imbalance leads to disparate visibility and exposure at the advantage of the most represented group. Recommendation effectiveness is decoupled from equity of visibility and exposure, with BPR returning the best trade-off between the two properties in the course-based representation.*

# 4 MITIGATING DISPARATE IMPACT

We mitigate disparities with a re-ranking algorithm that introduces items of the disadvantaged group in the recommendation list. A re-ranking is the only option when optimizing ranking-based metrics, like visibility and exposure. In-processing regularizations, such as [2, 14], would not be possible, since a model does not predict *if and where* an item will be ranked. Re-rankings have been employed to reduce disparities, both for non-personalized rankings [3, 7, 22, 28, 31, 32] and for recommender systems [5, 18, 20, 29], with approaches such as Maximal Marginal Relevance [6]. These optimize either visibility or exposure, so no comparison is possible.

## 4.1 Algorithm

The idea behind our mitigation algorithm is to *move up in the recommendation list the course that causes the minimum loss in prediction for all the learners.* Algorithm 1 describes the mitigation process, which is divided into three methods.

The first, $optimizeVisibilityExposure$ (lines 1-6), starts the mitigation. It makes two interventions: one based on visibility and the second one based on exposure. The second method, called *mitigation* (lines 7-29), regulates the visibility and exposure inside the recommendation list. The $checkPosition$ method (lines 30-34) is responsible for checking the position of an item in the list, taking into account if we perform a visibility- or exposure-based mitigation. The role of each line is commented in blue in the algorithm.

## 4.2 Impact of Mitigation

Tables 2 and 3 report the results after mitigating considering the course- and rating-based representations of the groups. Given the temporal split of the data, we cannot perform statistical tests to validate the results so, under each metric, we report the gain/loss obtained after running our mitigation. Our results present a general pattern, which leads us to our third observation.

**Observation 3**. *When providing a re-ranking based on minimal predicted loss, effectiveness remains stable, but disparate visibility and disparate exposure are mitigated. Interventions to adjust both visibility and exposure are needed to provide equity; if we mitigate only having a visibility goal, disparate exposure still occurs (fourth column, in red, in Tables 2 and 3).*

# 5 CONCLUSIONS AND FUTURE WORK

In this paper, we considered course recommender systems, with a focus on how teachers can be affected by the way courses are geographically distributed. Considering a real-world dataset coming from an online course platform, we assessed that the most represented country (the United States) is over-exposed by state-of-the-art recommendation models, affecting the teachers from the rest of the world. To overcome this issue, we proposed a re-ranking approach that aims to provide equity, by reaching the target visibility and exposure while causing the minimum loss in relevance.

In future work, we will go beyond this type of mitigation of the disparities, to re-distribute the recommendations in equitable ways between the individual countries, taking into account for multiple aspects (e.g., the language of the courses offered in non-English countries and that of the learners).

# ACKNOWLEDGMENTS

This research was partially funded by projects 2017-SGR-341, MISMIS-LANGUAGE (grant No. PGC2018-096212-B-C33) from the Spanish Ministry of Science and Innovation, NanoMoocs (grant No. COMRDI18-1-0010) from ACCIO, and PrEFair from ACCIO.

```
Input: recList: ranked list (records contain user, item, prediction, exposure, position)
Output: reRankedList: ranked list adjusted by visibility and exposure
define optimizeVisibilityExposure (recList)
begin
    reRankedList ← mitigation(recList, "visibility") ; // mitigation to target the desired visibility
    reRankedList ← mitigation(reRankedList, "exposure") ; // mitigation to regulate the exposure
    return reRankedList ; // return the re-ranked list
end
define mitigation (recList, reRankingType) // add the courses of the disadvantaged group to the top-k
begin
    for user ∈ list.users do // for each user
        for item ∈ top-n do // we loop over all items that belong to this user
            if checkPosition(item, itemsOut, reRankingType) is True then // check the position
                itemsOut.add(item) ; // add the item as possible candidate to move out to the list
            else if checkPosition(item, itemsOut, reRankingType) is False then
                itemsIn.add(item) ; // add the item as possible candidate to move in to the list
            end
        end
        while itemsIn not empty and itemsOut not empty do // computes all possible swaps and the loss of each one
            itemsIn ← itemsIn.pop(first); itemsOut ← itemsOut.pop(last); loss ← itemsOut.last − itemsIn.first;
            possibleSwaps.add(id,user,itemsOut.last,itemsIn.first,loss);
        end
    end
    if reRankingType == "visibility" then sortByLoss(possibleSwaps); // sort by loss in case of visibility ;
    else if reRankingType == "exposure" then sortByExposureLoss(possibleSwaps); // sort by exposure loss in case of exposure ;
    while proportions < targetProportions and possibleSwaps not empty do // do swaps until the target proportions are reached
        list ← swap(list, itemOut, itemIn); // makes the swap of the candidate with minor loss
        proportions ← updateProportions(itemOut, itemIn, reRankingType); // updates group proportions
    end
    return list; // returns the re-ranked list
end
define checkPosition(item, itemsOut, reRankingType) // check the position of an item in the list
begin
    if reRankingType == "visibility" then return item.position < top-k ;
    else if reRankingType == "exposure" then return item.position < itemsOut.last.position ;
end
```

**Algorithm 1:** Visibility and exposure mitigation algorithm.

**Table 2: Results for group $m$ of the mitigation based on $\mathcal{R}_C$, both after optimizing for Visibility and after optimizing for Exposure (here, we report only the NDCG and the disparate exposure; visibility, by design, remains the same).**

| | Visibility | | | Exposure | |
|---|---|---|---|---|---|
| **Algorithm** | **NDCG** | $\Delta\mathcal{V}_C$ | $\Delta\mathcal{E}_C$ | **NDCG** | $\Delta\mathcal{E}_C$ |
| **MostPop** | 0.0181 | 0.0000 | -0.0924 | 0.0166 | 0.0000 |
| *(gain/loss)* | -0.0012 | 0.3091 | 0.1193 | -0.0027 | 0.2117 |
| **RandomG** | 0.0006 | 0.0000 | -0.0001 | 0.0006 | 0.0000 |
| *(gain/loss)* | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **UserKNN** | 0.0369 | 0.0000 | -0.0233 | 0.0360 | 0.0000 |
| *(gain/loss)* | -0.0003 | 0.0402 | 0.1225 | -0.0012 | 0.1457 |
| **ItemKNN** | **0.2061** | 0.0000 | -0.0301 | **0.2038** | 0.0000 |
| *(gain/loss)* | -0.0008 | 0.0862 | 0.0481 | -0.0030 | 0.0782 |
| **BPR** | 0.1395 | 0.0000 | -0.0288 | 0.1375 | 0.0000 |
| *(gain/loss)* | -0.0006 | 0.0714 | 0.0370 | -0.0026 | 0.0658 |
| **BiasedMF** | 0.0007 | 0.0000 | -0.0266 | 0.0006 | 0.0000 |
| *(gain/loss)* | -0.0001 | 0.1064 | 0.0682 | -0.0001 | 0.0948 |
| **SVD++** | 0.0043 | 0.0000 | -0.0063 | 0.0043 | 0.0000 |
| *(gain/loss)* | -0.0001 | 0.0534 | 0.0480 | -0.0001 | 0.0543 |

**Table 3: Results for group $m$ of the mitigation based on $\mathcal{R}_R$, both after optimizing for Visibility and after optimizing for Exposure (here, we report only the NDCG and the disparate exposure; visibility, by design, remains the same).**

| | Visibility | | | Exposure | |
|---|---|---|---|---|---|
| **Algorithm** | **NDCG** | $\Delta\mathcal{V}_R$ | $\Delta\mathcal{E}_R$ | **NDCG** | $\Delta\mathcal{E}_R$ |
| **MostPop** | 0.0186 | 0.0000 | -0.0986 | 0.0178 | 0.0000 |
| *(gain/loss)* | -0.0007 | 0.2448 | 0.0488 | -0.0015 | 0.1474 |
| **RandomG** | 0.0006 | 0.0000 | 0.0217 | 0.0006 | 0.0000 |
| *(gain/loss)* | 0.0000 | -0.0644 | -0.0426 | 0.0000 | -0.0643 |
| **UserKNN** | 0.0369 | 0.0000 | -0.0237 | 0.0364 | 0.0000 |
| *(gain/loss)* | -0.0003 | -0.0241 | 0.0577 | -0.0008 | 0.0814 |
| **ItemKNN** | **0.2068** | 0.0000 | -0.0113 | **0.2061** | 0.0000 |
| *(gain/loss)* | 0.0000 | 0.0219 | 0.0026 | -0.0007 | 0.0139 |
| **BPR** | 0.1401 | 0.0000 | -0.0083 | 0.1396 | 0.0000 |
| *(gain/loss)* | 0.0000 | 0.0071 | -0.0069 | -0.0005 | 0.0015 |
| **BiasedMF** | 0.0007 | 0.0000 | -0.0060 | 0.0007 | 0.0000 |
| *(gain/loss)* | 0.0000 | 0.0421 | 0.0245 | 0.0000 | 0.0305 |
| **SVD++** | 0.0044 | 0.0000 | -0.0575 | 0.0045 | 0.0000 |
| *(gain/loss)* | 0.0000 | -0.0110 | -0.0675 | 0.0001 | -0.0101 |

## REFERENCES


[1] Alejandro Bellogín, Pablo Castells, and Iván Cantador. 2017. Statistical biases in Information Retrieval metrics for recommender systems. *Inf. Retr. Journal* 20, 6 (2017), 606–634.

[2] 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, KDD 2019*. ACM, 2212–2220. https://doi.org/10.1145/3292500.3330745

[3] 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, SIGIR 2018*. ACM, 405–414. https://doi.org/10.1145/3209978.3210063

[4] Ludovico Boratto, Gianni Fenu, and Mirko Marras. 2019. The Effect of Algorithmic Bias on Recommender Systems for Massive Open Online Courses. In *Advances in Information Retrieval - 41st European Conference on IR Research, ECIR 2019, Cologne, Germany, April 14-18, 2019, Proceedings, Part I (Lecture Notes in Computer Science, Vol. 11437)*, Leif Azzopardi, Benno Stein, Norbert Fuhr, Philipp Mayr, Claudia Hauff, and Djoerd Hiemstra (Eds.). Springer, 457–472. https://doi.org/10.1007/978-3-030-15712-8_30

[5] Robin Burke, Nasim Sonboli, and Aldo Ordonez-Gauger. 2018. Balanced Neighborhoods for Multi-sided Fairness in Recommendation. In *Conference on Fairness, Accountability and Transparency, FAT 2018 (Proceedings of Machine Learning Research, Vol. 81)*. PMLR, 202–214. http://proceedings.mlr.press/v81/burke18a.html

[6] Jaime G. Carbonell and Jade Goldstein. 1998. The Use of MMR, Diversity-Based Reranking for Reordering Documents and Producing Summaries. In *SIGIR '98: Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 335–336. https://doi.org/10.1145/290941.291025

[7] 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 (LIPIcs, Vol. 107)*. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 28:1–28:15. https://doi.org/10.4230/LIPIcs.ICALP.2018.28

[8] Paolo Cremonesi, Yehuda Koren, and Roberto Turrin. 2010. Performance of recommender algorithms on top-n recommendation tasks. In *Proceedings of the 2010 ACM Conference on Recommender Systems, RecSys 2010*. ACM, 39–46. https://doi.org/10.1145/1864708.1864721

[9] Danilo Dessì, Gianni Fenu, Mirko Marras, and Diego Reforgiato Recupero. 2018. COCO: Semantic-Enriched Collection of Online Courses at Scale with Experimental Use Cases. In *Trends and Advances in Information Systems and Technologies - Volume 2 [WorldCIST'18, Naples, Italy, March 27-29, 2018] (Advances in Intelligent Systems and Computing, Vol. 746)*, Álvaro Rocha, Hojjat Adeli, Luís Paulo Reis, and Sandra Costanzo (Eds.). Springer, 1386–1396. https://doi.org/10.1007/978-3-319-77712-2_133

[10] Francesco Fabbri, Francesco Bonchi, Ludovico Boratto, and Carlos Castillo. 2020. The Effect of Homophily on Disparate Visibility of Minorities in People Recommender Systems. In *Proceedings of the Fourteenth International AAAI Conference on Web and Social Media, ICWSM 2020*. AAAI Press, 165–175. https://aaai.org/ojs/index.php/ICWSM/article/view/7288

[11] Sara Hajian, Francesco Bonchi, and Carlos Castillo. 2016. Algorithmic Bias: From Discrimination Discovery to Fairness-aware Data Mining. In *ACM SIGKDD International Conf. on Knowledge Discovery and Data Mining*. ACM, 2125–2126.

[12] Jonathan L. Herlocker, Joseph A. Konstan, and John Riedl. 2002. An Empirical Analysis of Design Choices in Neighborhood-Based Collaborative Filtering Algorithms. *Inf. Retr.* 5, 4 (2002), 287–310. https://doi.org/10.1023/A:1020443909834

[13] Kalervo Järvelin and Jaana Kekäläinen. 2002. Cumulated gain-based evaluation of IR techniques. *ACM Trans. Inf. Syst.* 20, 4 (2002), 422–446. https://doi.org/10.1145/582415.582418

[14] Toshihiro Kamishima, Shotaro Akaho, Hideki Asoh, and Jun Sakuma. 2018. Recommendation Independence. In *Conference on Fairness, Accountability and Transparency, FAT 2018 (Proceedings of Machine Learning Research, Vol. 81)*. PMLR, 187–201. http://proceedings.mlr.press/v81/kamishima18a.html

[15] Yeqin Kang. 2014. An analysis on SPOC: Post-MOOC era of online education. *Tsinghua Journal of Education* 35, 1 (2014), 85–93.

[16] Yehuda Koren. 2008. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 426–434. https://doi.org/10.1145/1401890.1401944

[17] Yehuda Koren, Robert M. Bell, and Chris Volinsky. 2009. Matrix Factorization Techniques for Recommender Systems. *IEEE Computer* 42, 8 (2009), 30–37. https://doi.org/10.1109/MC.2009.263

[18] Weiwen Liu, Jun Guo, Nasim Sonboli, Robin Burke, and Shengyu Zhang. 2019. Personalized fairness-aware re-ranking for microlending. 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, 467–471. https://doi.org/10.1145/3298689.3347016

[19] Mirko Marras, Ludovico Boratto, Guilherme Ramos, and Gianni Fenu. 2020. Equality of Learning Opportunity via Individual Fairness in Personalized Recommendations. *CoRR* abs/2006.04282 (2020). arXiv:2006.04282 https://arxiv.org/abs/2006.04282

[20] Rishabh Mehrotra, James McInerney, Hugues Bouchard, Mounia Lalmas, and Fernando Diaz. 2018. Towards a Fair Marketplace: Counterfactual Evaluation of the trade-off between Relevance, Fairness & Satisfaction in Recommendation Systems. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management, CIKM 2018*. ACM, 2243–2251. https://doi.org/10.1145/3269206.3272027

[21] Xia Ning, Christian Desrosiers, and George Karypis. 2015. A Comprehensive Survey of Neighborhood-Based Recommendation Methods. In *Recommender Systems Handbook*, Francesco Ricci, Lior Rokach, and Bracha Shapira (Eds.). Springer, 37–76. https://doi.org/10.1007/978-1-4899-7637-6_2

[22] Gourab K. Patro, Arpita Biswas, Niloy Ganguly, Krishna P. Gummadi, and Abhijnan Chakraborty. 2020. FairRec: Two-Sided Fairness for Personalized Recommendations in Two-Sided Platforms. In *WWW '20: The Web Conference 2020*. ACM / IW3C2, 1194–1204. https://doi.org/10.1145/3366423.3380196

[23] Guilherme Ramos and Ludovico Boratto. 2020. Reputation (In)dependence in Ranking Systems: Demographics Influence Over Output Disparities. 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, 2061–2064. https://doi.org/10.1145/3397271.3401278

[24] Guilherme Ramos, Ludovico Boratto, and Carlos Caleiro. 2020. On the negative impact of social influence in recommender systems: A study of bribery in collaborative hybrid algorithms. *Inf. Process. Manag.* 57, 2 (2020), 102058. https://doi.org/10.1016/j.ipm.2019.102058

[25] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian Personalized Ranking from Implicit Feedback. In *UAI 2009, Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*. AUAI Press, 452–461. https://dslpitt.org/uai/displayArticleDetails.jsp?mmnu=1&smnu=2&article_id=1630&proceeding_id=25

[26] Badrul Munir Sarwar, George Karypis, Joseph A. Konstan, and John Riedl. 2001. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the Tenth International World Wide Web Conference, WWW 10*. ACM, 285–295. https://doi.org/10.1145/371920.372071

[27] João Saúde, Guilherme Ramos, Carlos Caleiro, and Soummya Kar. 2017. Reputation-Based Ranking Systems and Their Resistance to Bribery. In *2017 IEEE International Conference on Data Mining, ICDM 2017, New Orleans, LA, USA, November 18-21, 2017*, Vijay Raghavan, Srinivas Aluru, George Karypis, Lucio Miele, and Xindong Wu (Eds.). IEEE Computer Society, 1063–1068. https://doi.org/10.1109/ICDM.2017.139

[28] 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, KDD 2018*. ACM, 2219–2228. https://doi.org/10.1145/3219819.3220088

[29] Nasim Sonboli, Farzad Eskandanian, Robin Burke, Weiwen Liu, and Bamshad Mobasher. 2020. Opportunistic Multi-aspect Fairness through Personalized Re-ranking. 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, 239–247. https://doi.org/10.1145/3340631.3394846

[30] Elaine Walster, Ellen Berscheid, and G William Walster. 1973. New directions in equity research. *Journal of personality and social psychology* 25, 2 (1973), 151.

[31] 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, CIKM 2017*. ACM, 1569–1578. https://doi.org/10.1145/3132847.3132938

[32] Meike Zehlike and Carlos Castillo. 2020. Reducing Disparate Exposure in Ranking: A Learning To Rank Approach. In *WWW '20: The Web Conference 2020*. ACM / IW3C2, 2849–2855. https://doi.org/10.1145/3366424.3380048