Bias in Book Recommendation

Savvina Daniil
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands
s.daniil@cwi.nl

ABSTRACT

Recommender systems are prevalent in many applications, but hide risks; issues like bias propagation have been on the focus of related studies in recent years. My own research revolves around tracking bias in the book recommendation domain. Specifically, I am interested in whether the incorporation of recommender systems in a library's loaning system serves their social responsibility and purpose, with bias being the main point of concern. To this end, I engage with the topic in three ways; by mapping the area of ethics in book recommendation, by investigating and reflecting on challenges with studying bias in recommender systems in general, and by showcasing a set of social implication of statistical bias in the book recommendation domain in particular. In this doctoral symposium paper, I further elaborate on the problem at hand, the outline of my thesis, the progress I have made so far, as well as my plans for future work along with specific questions that have arisen from my research efforts.

CCS CONCEPTS

• **Information systems** → **Recommender systems**; • **Applied computing** → *Arts and humanities*.

KEYWORDS

Recommender Systems, Bias, Books, Reproducibility

ACM Reference Format:
Savvina Daniil. 2024. Bias in Book Recommendation. In *18th ACM Conference on Recommender Systems (RecSys '24), October 14–18, 2024, Bari, Italy*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3640457.3688025

1 INTRODUCTION

Bias in recommender systems is often defined and studied in two forms: either as statistical bias that certain algorithms are prone to (e.g., popularity bias) or as social bias that exists in some form in the data and results to unfairness (e.g., a group of people is underrepresented) [2]. These forms of bias are not unrelated but rather interconnected; social bias in the data can interact with statistical bias on the side of the algorithms and persist in the recommendations that users receive. It is, therefore, relevant to study the link between them in the context of book recommender systems, which is the main topic of my thesis.

RecSys '24, October 14–18, 2024, Bari, Italy

ACM ISBN 979-8-4007-0505-2/24/10
https://doi.org/10.1145/3640457.3688025

Studying statistical bias in recommender systems is generally challenging. Recommender systems as a discipline is known to suffer from reproducibility issues that relate to lack of appropriate documentation on behalf of research studies, as well as the inherent application-specific nature of such systems [6]. In my thesis, I will discuss and experimentally investigate the effect of different methodological decisions on popularity bias.

The implications of data bias heavily depend on the domain at hand. Studies often focus on breaking down the ethical issues of recommender systems into multiple perspectives depending on the stakeholders of a system [11]. In the case of book recommender systems bias, one can borrow from the field of Library and Information Science (LIS) to map these perspectives. Mathiesen [10] introduces a framework to describe informational justice in library services with the following components: seekers (users), sources (authors), and subjects (characters/themes). In my thesis, I will focus on book authors, who are known to suffer from social bias in book publishing in general [7, 14].

My research focuses on three aspects of bias in book recommender systems. First, I am mapping the current ethical considerations concerning book recommender systems, while zooming in on the topic of bias, and to a lesser extent, diversity. The goal is to provide a good overview of the field and track what is the current state and open issues. Second, I am taking a critical look into the practice of measuring and reporting on statistical bias in recommendation in order to raise awareness with regards to the challenges present in research. Third, I am investigating the implications of statistical bias from a social perspective, specifically when it comes to book authors, as a means to highlight the interconnection between statistical and social bias in the book recommendation domain. In the following sections, I will outline my thesis, reflect on the progress I've made so far, and describe the future work ahead. Through this doctoral symposium paper, I hope to get input and feedback from the scholar community at RecSys.

2 THESIS OUTLINE

In this section I will introduce the scope of my PhD thesis, list challenges that I have faced, and outline the planned content and structure.

2.1 Scope and challenges

The topic of bias in book recommender systems can be approached from different perspectives. What is specifically in scope for my thesis are the following:

- What are the ethical issues around book recommender systems? What perspectives currently exist?
- What are the challenges of studying bias in recommender systems? What specific characteristics of a research process

can have an impact on whether bias will be observed as a result of a recommendation?
- What are the societal implications of statistical biases in book recommender systems? How can statistical biases in a dataset of user-item interactions translate into social bias towards book authors?

In the process of answering these questions, I have come across a set of challenges. Due to lack of access to real world data, my research is limited to the few benchmark datasets that are available online, as well as synthetic data. Additionally, the topic of bias in book recommendation has received little attention from the research community compared to other media. The challenges of limited datasets and limited research are interconnected. As a result, part of my work focused on proposing ways to synthesize data, as well as enriching existing data with information on the authors included.

2.2 Content

To outline my thesis, I have split it into three parts that correspond to the relevant aspects described in the previous section. In this section, I describe the chapters that will correspond to each part.

2.2.1 Part 1: Understanding ethical issues of book recommender systems.

Chapter 1: A survey on bias in book recommendation. While bias in recommendation has received significant attention by the research community, this is not the case for bias in book recommendation specifically. Experimental research seems to focus more on other types of media recommendation, with a few exceptions [5, 12]. There is theoretical work laid out from LIS that attempts to map ethical issues of information access, but experimental work often does not build on it. In this chapter, I will review the relevant work, both theoretical and experimental, map the area of ethics in book recommender systems, and present developments, challenges, and open questions. Such a survey is currently lacking in literature, and it will also allow me to build from there by zooming in on specific ethical concerns.

Chapter 2: The perspective of public media service practitioners on diversity in recommender systems. In this chapter, I will zoom in on a subfield of ethics in recommender systems, namely diversity. In the same line as the previous chapter, it seems that the normative background is often missing from quantitative research into diversity in media recommendation in general. The perspective of practitioners on the matter of normative values and how they should be incorporated in a system is often not taken into account. Interviewing them about the importance and operationalization of diversity can offer very interesting insights.

2.2.2 Part 2: A critical look into measuring statistical bias in book recommender systems.

Chapter 3: Reproducing bias in recommendation studies. Recommender systems as a field are known to suffer from a reproducibility problem. Research is often fragmented and domain-specific. This is also the case for bias in recommender systems. In this chapter, I will reproduce three studies on the topic of popularity bias in recommendation and attempt to uncover why their results differ despite claims of following a similar approach.

Chapter 4: The challenges of studying bias in recommender systems. The topic of which aspects of the recommendation process can impact whether bias propagation is observed is very relevant for the research community. The process of showcasing the dependence of reporting on details can extend further from reproducing existing research. By controlling and experimenting with relevant aspects, we can observe which ones influence bias and therefore study it in a systematic manner. In this chapter, I will experiment with combining synthetic datasets of user-item interactions and different versions of the same recommender systems algorithms to observe their combined effect on popularity bias.

2.2.3 Part 3: Societal implications of statistical bias in book recommender systems.

Chapter 5: Author bias in book recommendation. In the final chapter, I will focus on bias in book recommendation from the perspective of authors. To do so, I will enrich a set of available datasets with user-book interactions with information about the author that is publicly available. Afterwards, I will combine the datasets with various recommender systems algorithms and observe whether statistical biases also lead to social biases towards the authors individual characteristics. The results from this study will form the final chapter of my thesis.

3 PROGRESS SO FAR

In this section, I will present the studies I have carried out so far as a part of my PhD. They are organized in chronological order and not based on the order that they will be placed in the thesis itself.

3.1 Hidden Author Bias in Book Recommendation

This study will become the basis of the fifth chapter of my thesis. It was presented at the FAccTRec workshop of ACM RecSys 2022. We plan to extend it into a full paper.

In this study [3], we answered the following research questions:

(1) Do commonly used recommender algorithms propagate data bias towards author country of citizenship?
(2) What is the relation between author country of citizen bias and popularity bias?

We studied these questions on a commonly used dataset with book ratings, Book-Crossing [18]. We enriched it with author information using publicly available external sources, as shown in figure 1. We found that popular books are mainly written by US citizens in the dataset, and that these books tend to be recommended disproportionally by popular collaborative filtering algorithms compared to the users' profiles, as shown in figures 2 and 3.

The results show that social bias in book recommender systems can occur as a direct result of popularity bias.

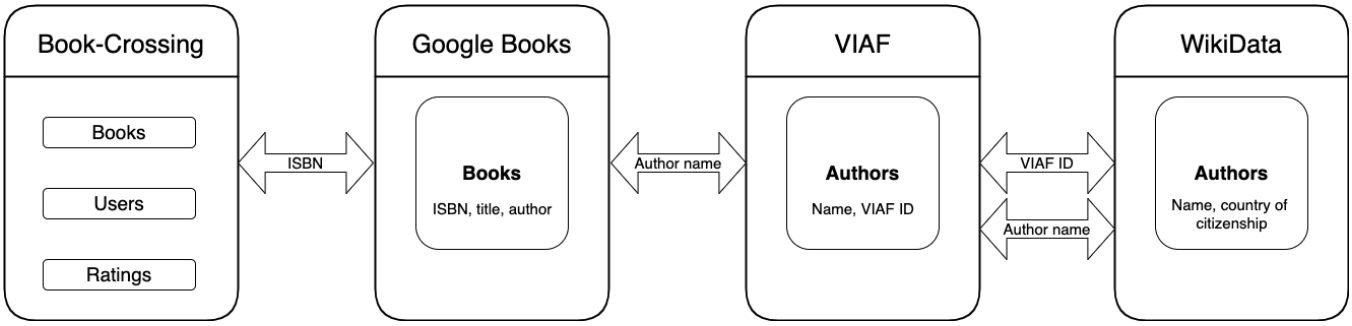

Figure 1: Links between the datasets.

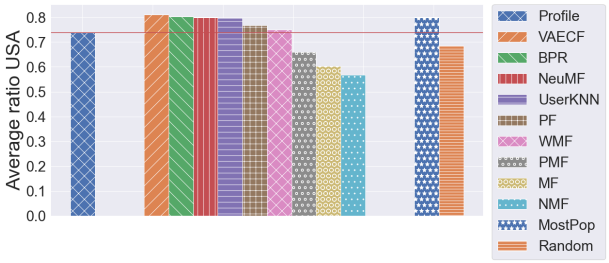

Figure 2: Average ratio of recommended books by every algorithm that were written by US citizens. Comparison with the average ratio of American-authored books in the users' profiles.

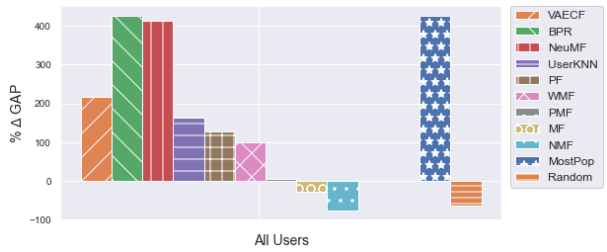

Figure 3: Relative increase in average popularity between profile and recommendation by every algorithm, averaged over all users.

3.2 Reproducing Popularity Bias in Recommendation: The Effect of Evaluation Strategies

This study will be the third chapter of my thesis. It was published in the special issue on Perspectives on Recommender Systems Evaluation of the ACM journal Transactions On Recommender Systems (ACM TORS).

Recent work focused on the topic of uneven popularity bias propagation among users with varying interests for niche items, with movies being the domain of interest [1]. Later on, two different research teams reproduced the methodology in the domains of music [9] and books [12] respectively. The results across the different domains diverge. In this paper [4], we reproduced the three studies and identify four aspects that are relevant in investigating the differences in results: data, algorithms, division of users in groups and evaluation strategy. We ran a set of experiments in which we measured general popularity bias propagation and unfair treatment of certain users with various combinations of these aspects. We found that all aspects account to some degree for the divergence in results, and should be carefully considered in future studies. Importantly, we found that the divergence in findings can be in large part attributed to the choice of evaluation strategy.

See figure 4 for an example of the impact of each aspect. Each row corresponds to a dataset: MovieLens1M [8], LastFM [16], or Book-Crossing [12, 18]. Each column corresponds to an evaluation strategy, with the three strategies differing between them in terms of the user and item candidates [15]. Figure 4 shows that whether item popularity in profile and recommendation frequency correlate highly depends on the combination of aspects.

3.3 Diversity of What? On the Different Conceptualizations of Diversity in Recommender Systems

This study will be the second chapter of my thesis. It was accepted and presented at ACM FAccT 2024.

Diversity is a commonly known principle in the design of recommender systems, but also ambiguous in its conceptualization. In this study [17], through semi-structured interviews we explored how practitioners at three different public service media organizations in the Netherlands, including a library, conceptualize diversity within the scope of their recommender systems. We provided an overview of the goals that they have with diversity in their systems, which aspects are relevant (as seen in figure 5), and how recommendations should be diversified. We showed that even within this limited domain, conceptualization of diversity greatly varies, and argue that it is unlikely that a standardized conceptualization will be achieved. Instead, we should focus on effective communication of what diversity in this particular system means, thus allowing for operationalizations of diversity that are capable of expressing the nuances and requirements of that particular domain.

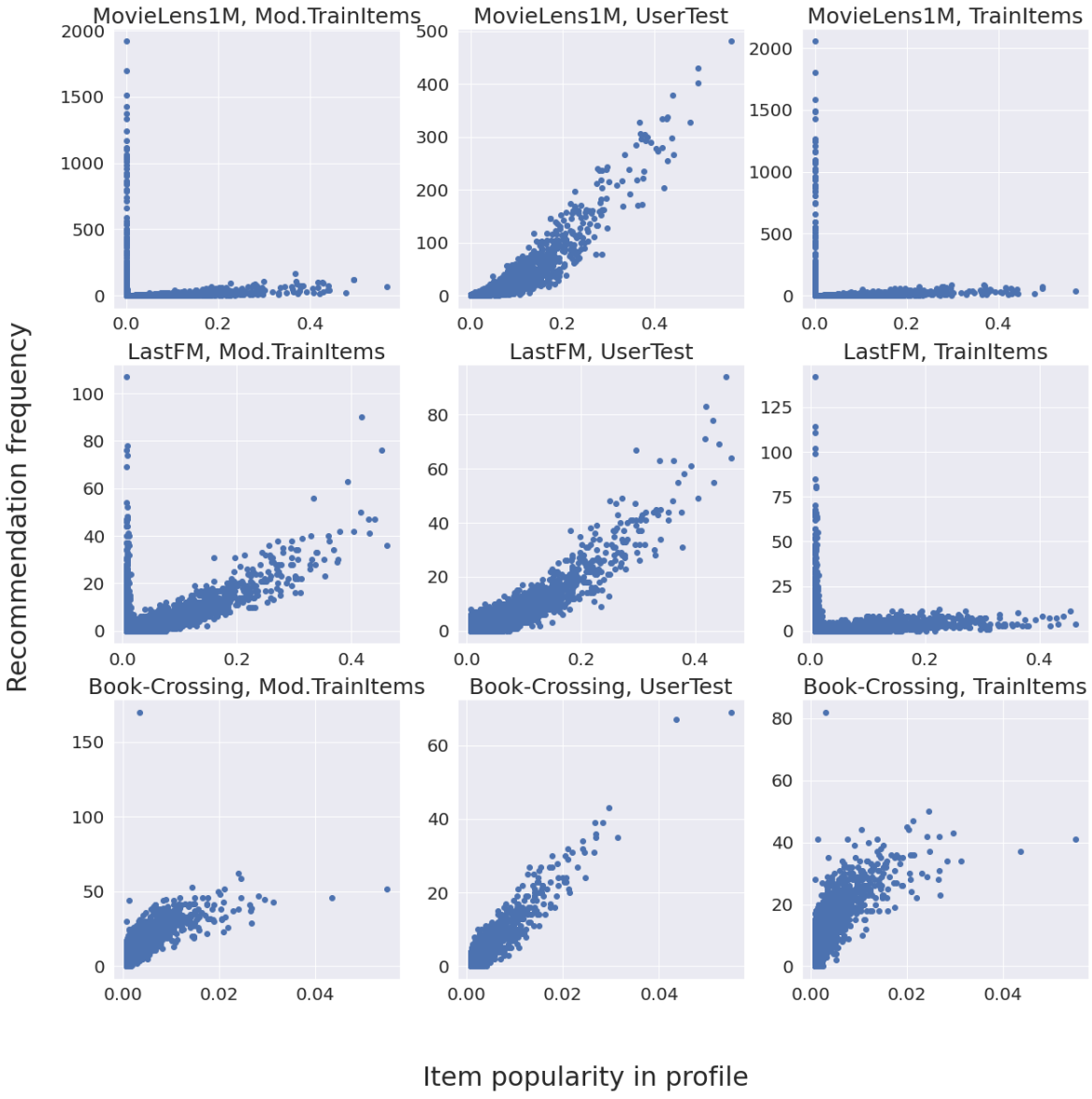

Figure 4: Item popularity in profile versus frequency of recommendation by the algorithm ItemKNN, for every dataset and evaluation strategy tested.

3.4 On the Challenges of Studying Bias in Recommender Systems: The Effect of Data Characteristics and Algorithm Configuration

This study will be the fourth chapter of my thesis.

In this study, we explored the challenges of measuring and reporting popularity bias. We showcased the impact of data properties and algorithm configurations on popularity bias by combining real and synthetic data with well known recommender systems frameworks. First, we identified data characteristics that might impact popularity bias. Accordingly, we generated various datasets that combine

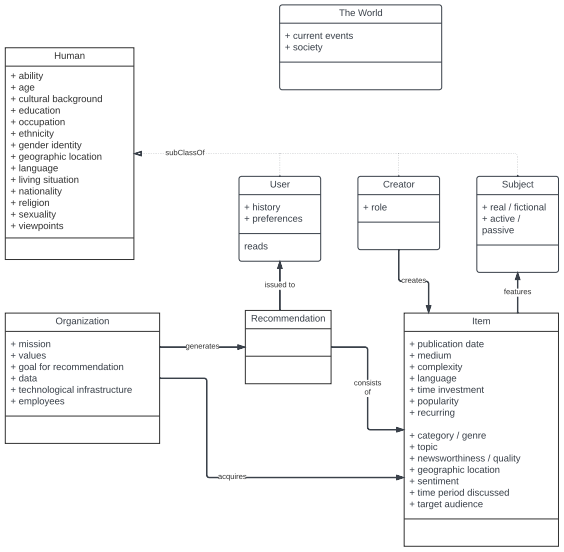

Figure 5: Schematic overview of the identified aspects of diversity and how they interact with each other.

these characteristics, based on a dataset of interactions between users and books. Second, we located algorithm configurations that vary across implementations in literature. We evaluated popularity bias for a number of datasets, three real and five synthetic, and configurations, and offered insights on their joint effect. We found that, depending on the data characteristics, various configurations of the algorithms examined can lead to different conclusions regarding the propagation of popularity bias. These results motivate the need for explicitly addressing algorithmic configuration and data properties when reporting and interpreting bias in recommender systems.

4 FUTURE WORK

In this section, I will outline the plans for the missing chapters.

4.1 Chapter 1: A survey in bias in book recommendation

For this chapter, I plan to perform an extensive survey on ethics in book recommendation, with a focus on bias, both from a theoretical and a practical perspective. I have organized a set of keywords to search for, and plan to use ideas from the PRISMA methodology [13] to guide the survey and enhance transparency. Due to limited experimental work on the topic, the survey may be relatively small.

4.2 Chapter 4: Studying bias in recommender systems

While I have extensively worked on this and there is a draft, as described in the previous section, getting this study published has proven difficult, despite reviewers generally agreeing that the ideas are valuable. One reason seems to be lack of focus on state-of-the-art graph and/or neural network based approaches. I would be very interested in receiving feedback from the community on what additions and corrections would be needed for this study to be considered sufficiently refined.

4.3 Chapter 5: Author bias in book recommendation

For the final chapter of my thesis, my intention is to expand on the paper presented at FAccTRec 2022 (see section 3.1) by following the same process but for other datasets. As such, I will perform and present a full analysis on the implications of popularity bias for book authors for various datasets available online. One of the datasets that I will include in the experimentation was made available to me by the Danish public libraries.

5 CONCLUSION

In this paper, I introduced the topic of bias in book recommender systems and highlighted three essential components of studying it. I specified what is in scope in the context of my PhD thesis, as well as challenges that I have faced in the process of performing research. I outlined the content of my thesis that shall emerge from my research, chapter by chapter. I described my progress by going through the studies I have carried out so far and linking them to the appropriate chapters. Finally, I laid out my plan for future work and addressed specific questions to the scholarly community at RecSys.

REFERENCES

[1] Himan Abdollahpouri, Masoud Mansoury, Robin Burke, and Bamshad Mobasher. 2019. The Unfairness of Popularity Bias in Recommendation. In *Recommendation in Multi-stakeholder Environments (RMSE), in conjunction with the 13th ACM Conference on Recommender Systems*.

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

[3] Savvina Daniil, Mirjam Cuper, Cynthia Liem, Jacco van Ossenbruggen, and Laura Hollink. 2022. Hidden author bias in book recommendation. *arXiv preprint arXiv:2209.00371* (2022).

[4] Savvina Daniil, Mirjam Cuper, Cynthia CS Liem, Jacco van Ossenbruggen, and Laura Hollink. 2024. Reproducing popularity bias in recommendation: The effect of evaluation strategies. *ACM Transactions on Recommender Systems* 2, 1 (2024), 1–39.

[5] Michael D Ekstrand and Joseph A Konstan. 2019. Recommender systems notation: proposed common notation for teaching and research. *arXiv preprint arXiv:1902.01348* (2019).

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

[7] Nettie Finn. 2016. Pseudonymous Disguises: Are Pen Names An Escape From the Gender Bias in Publishing? (2016).

[8] F Maxwell Harper and Joseph A Konstan. 2015. The movielens datasets: History and context. *Acm transactions on interactive intelligent systems (tiis)* 5, 4 (2015), 1–19.

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

[10] Kay Mathiesen. 2015. Informational justice: A conceptual framework for social justice in library and information services. *Library trends* 64, 2 (2015), 198–225.

[11] Silvia Milano, Mariarosaria Taddeo, and Luciano Floridi. 2021. Ethical aspects of multi-stakeholder recommendation systems. *The information society* 37, 1 (2021), 35–45.

[12] Mohammadmehdi Naghiaei, Hossein A Rahmani, and Mahdi Dehghan. 2022. The unfairness of popularity bias in book recommendation. In *Advances in Bias and Fairness in Information Retrieval: Third International Workshop, BIAS 2022, Stavanger, Norway, Revised Selected Papers*. Springer, 69–81.

[13] Matthew J Page, Joanne E McKenzie, Patrick M Bossuyt, Isabelle Boutron, Tammy C Hoffmann, Cynthia D Mulrow, Larissa Shamseer, Jennifer M Tetzlaff, Elie A Akl, Sue E Brennan, et al. 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Bmj* 372 (2021).

[14] Anamik Saha and Sandra Van Lente. 2020. Rethinking'Diversity'in Publishing. (2020).

[15] Alan Said and Alejandro Bellogín. 2014. Comparative recommender system evaluation: benchmarking recommendation frameworks. In *Proceedings of the 8th ACM Conference on Recommender systems*. 129–136.

[16] Markus Schedl. 2016. The LFM-1b Dataset for Music Retrieval and Recommendation. In *Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval* (New York, New York, USA) *(ICMR '16)*. Association for Computing Machinery, New York, NY, USA, 103–110. https://doi.org/10.1145/2911996.2912004

[17] Sanne Vrijenhoek, Savvina Daniil, Jorden Sandel, and Laura Hollink. 2024. Diversity of What? On the Different Conceptualizations of Diversity in Recommender Systems. *arXiv preprint arXiv:2405.02026* (2024).

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