Retrieval and Recommendation Systems at the Crossroads of Artificial Intelligence, Ethics, and Regulation

Markus Schedl
Johannes Kepler University and Linz Institue of Technology
Linz, Austria
markus.schedl@jku.at

Emilia Gómez
European Commission, Joint Research Centre and Universitat Pompeu Fabra
Seville/Barcelona, Spain
emilia.gomez-gutierrez@ec.europa.eu

Elisabeth Lex
Graz University of Technology
Graz, Austria
elisabeth.lex@tugraz.at

ABSTRACT

This tutorial aims at providing its audience an interdisciplinary overview about the topics of fairness and non-discrimination, diversity, and transparency of AI systems, tailored to the research fields of information retrieval and recommender systems. By means of this tutorial, we would like to equip the mostly technical audience of SIGIR with the necessary understanding of the ethical implications of their research and development on the one hand, and of recent political and legal regulations that address the aforementioned challenges on the other hand.

CCS CONCEPTS

• **Information systems → Recommender systems**; **Document filtering**; • **Applied computing** → *Law, social and behavioral sciences.*

KEYWORDS

recommender systems, information retrieval, ethics, fairness, non-discrimination, diversity, transparency, regulation

ACM Reference Format:
Markus Schedl, Emilia Gómez, and Elisabeth Lex. 2022. Retrieval and Recommendation Systems at the Crossroads of Artificial Intelligence, Ethics, and Regulation. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '22), July 11–15, 2022, Madrid, Spain.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3477495.3532683

COVER SHEET

Duration: 3 hours plus breaks
Tutorial format: on-site event
Intended audience: The interdisciplinary tutorial addresses an intermediate audience in terms of information retrieval and recommender systems expertise. Since the main audience of SIGIR has a technical background, we do not assume knowledge in the other disciplines the tutorial connects to, i.e., policy, ethics, or law.

SIGIR '22, July 11–15, 2022, Madrid, Spain

ACM ISBN 978-1-4503-8732-3/22/07…$15.00
https://doi.org/10.1145/3477495.3532683

Contact information of presenters:
Markus Schedl: markus.schedl@jku.at (main contact);
Emilia Gómez: emilia.gomez-gutierrez@ec.europa.eu;
Elisabeth Lex: elisabeth.lex@tugraz.at

Brief biography of presenters:
Markus Schedl (http://www.mschedl.eu) is a full professor at the Johannes Kepler University Linz (JKU), affiliated with the Institute of Computational Perception, leading the Multimedia Mining and Search group. In addition, he is head of the Human-centered AI group at the Linz Institute of Technology (LIT) AI Lab. His main research interests include recommender systems, user modeling, information retrieval, machine learning, multimedia processing, and trustworthy AI, with a particular focus on detecting and mitigating bias in retrieval and recommendation algorithms [21, 25, 26, 35] and on psychological models for recommendation [22, 23, 34]. He (co-)authored more than 240 refereed conference papers, journal articles, and book chapters. He has already given numerous tutorials in top venues including ACM SIGIR (2013 on "Music Similarity and Retrieval" and 2015 on "Music Retrieval and Recommendation"), ACM Recommender Systems (2018 on "New Paths in Music Recommender Systems Research"), ACM Multimedia (2013 on "Multimedia Information Retrieval: Music and Audio"), and the World Wide Web conference (2018 on "Complex Recommendations" and 2022 on "Psychology-informed Recommender Systems: A Human-centric Perspective on Recommender Systems"). In addition, he has more than 15 years of experience as a lecturer at various national and international universities. He has recently co-authored an article about the topic of the tutorial, published in the Communications of the ACM [9].

Emilia Gómez (https://emiliagomez.com) holds BSc and MSc degrees in Electrical Engineering and a PhD degree in Computer Science. She is a principal investigator on Human and Machine Intelligence (HUMAINT) at the Joint Research Centre (European Commission). She is also a guest professor at the Music Technology Group, Universitat Pompeu Fabra, Barcelona. Her research is grounded in the Music Information Retrieval field, where she has developed data-driven technologies to support music listening experiences. Starting from music, she studies the impact of artificial intelligence (AI) on human decision making, cognitive and socio-emotional development. Her research interests include fairness and transparency in AI, the impact of AI on jobs, and how it affects children development. She is currently a member of the Spanish National Council for AI and the OECD One AI expert group.

Elisabeth Lex (https://elisabethlex.info) is an associate professor and principal investigator of the Recommender Systems and Social Computing Lab at Graz University of Technology (TUG). Her research interests include recommender systems, user modeling, information retrieval and computational social science, with a particular focus on psychology-informed recommender systems [15, 16, 18, 22, 23, 34, 39], bias in recommender systems [19, 21], human decision making and recommender systems [4, 8], privacy in recommender systems [29], or music consumption [17, 37]. Elisabeth has (co-)authored more than 120 peer-reviewed publications in the aforementioned topics. She has given tutorials on "Psychology-informed Recommender Systems" at the 11th Italian Information Retrieval Workshop (IIR) 2021, at the Complex Networks and their Application conference 2021, at the 7th ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR) 2022, and at The ACM World Wide Web conference 2022.

EXTENDED ABSTRACT

Motivation

Information retrieval (IR) recommender systems (RSs) affect many aspects of our daily lives, deciding which content we are exposed to on the web or social media platforms, which products to buy, or which music to listen to. With the ever increasing adoption of — mostly opaque — machine and deep learning technology in such systems, many ethical questions about their use have emerged. In particular, questions related to *fairness*, *non-discrimination*, *diversity*, and *transparency* have recently been in the focus of the public debate as well as discussed in many recent research articles, e.g., [5, 9, 10]. Therefore, we address those in the tutorial, and discuss them from an interdisciplinary point of view.

Fairness and Non-Discrimination. The discussion has been fueled by findings of recent studies that identified harmful biases in data, algorithmic behavior, and corresponding lists of retrieved documents and recommended items, e.g., [6, 14, 20, 21, 24, 35, 47]. These biases can result in unfair treatment or even discrimination against certain users or groups of users, e.g., with respect to their gender [20], age [38], or personality traits [26]. In some, but not all, cases such algorithmic behavior is illegal [10, 45].

Diversity. Studies have shown the value of diversity to improve innovation and excellence in research [42]. In the context of artificial intelligence (AI), several policy reports and experts [13, 44] have suggested as well to incorporate diversity in the AI development process. Diversity refers to the existence of variations of different characteristics among individuals, such as gender, age, race, religion, or cultural background, being related to the fairness principle mentioned above. AI systems, among which retrieval and recommender systems play a major role, should then incorporate a diversity of perspectives in research and development (e.g., through diverse research communities [12], developing teams or user groups) and make sure that developed technology provides an equal outcome for all potential stakeholders. Note that this does not only apply to the research communities and development teams, but in an information retrieval and recommender systems context also to content producers (e.g., diversity of authors of web documents that are retrieved, or music artists whose songs are recommended).

Transparency. Transparency has been defined as a means for trust in technologies and involves different concepts such as explainability, traceability, and communication [13, 40, 41, 46]. Explainability concerns the ability to explain the technical process of an AI system (i.e., provide the means for humans to understand and trace the outputs of the system) and the related human decisions (e.g., application domain or task to be solved), e.g., [30, 43]. These explanations should be adapted to different expertise levels, from developers to end users of the system. The related concept of justification refers to the requirement of a retrieval or recommendation system, in our case, to justify why a certain document or item was presented to the user, e.g., [1, 7]. Traceability allows keeping track of the behavior of a system in a chronological way, and facilitates auditability, i.e., the ethical assessment of algorithms to investigate potentially harmful consequences such as if an algorithm is biased or exhibits discriminatory behavior [3]. For selected works on auditing algorithms please refer to, e.g., [2, 27, 33, 36]. Finally, the concept of communication incorporates the idea of documenting the system development process, capabilities, and limitations [28, 32].

The importance of these topics is further highlighted by many recent guidelines, regulations, and policies such as the ones in the EU and US, as discussed in [9, 31]. For instance, in the EU context, we can rely on the *EU Charter of Fundamental Rights*[1] [11], *EU Ethical Principles for Trustworthy AI*[2] [13], *Regulatory Framework for AI*,[3] and the *Digital Service Act*[4], which all strongly refer to retrieval and recommendation systems. In the US context, the *Platform Accountability and Transparency Act* (PATA),[5] proposed by several US senators, requires large platforms to make data available to support scientific research and oversight connected to data-driven algorithms.

Since the topics of fairness, non-discrimination, diversity, and transparency affect the entire population and are influenced by many stakeholders, e.g., researchers, developers, policy makers, and economists, they call for an interdisciplinary treatment, involving the disciplines of artificial intelligence, computer science, ethics, legal, and political aspects, just to mention a few. Acknowledging these facts, the tutorial takes an interdisciplinary approach. Nevertheless, we particularly tailor our discussion of these topics to the SIGIR community. This means we consider information access systems, more precisely information retrieval and recommender systems.

Objectives

This tutorial aims at providing its audience an interdisciplinary overview about the topics of fairness and non-discrimination, diversity, and transparency of AI systems, tailored to the research fields of information retrieval and recommender systems. By means of this tutorial, we would like to equip the mostly technical audience

[1]https://ec.europa.eu/info/aid-development-cooperation-fundamental-rights/your-rights-eu/eu-charter-fundamental-rights_en
[2]https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1
[3]https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai
[4]https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package
[5]http://www.coons.senate.gov/download/text-pata-117

of SIGIR with the necessary understanding of the ethical implications of their research and development on the one hand, and of recent political and legal regulations that address the aforementioned challenges on the other hand. As for these political and legal regulations, the tutorial foremost takes a European perspective, since EU regulation is at the forefront of elaborating guidelines for ethical and trustworthy AI (see previous section). Nevertheless, we also briefly review initiatives outside of Europe, in particular in the US.

Since the addressed topics are vital and relevant on a global scale, we strongly believe that the tutorial attracts a global audience, too. In particular, research in information retrieval and recommender systems has become a global endeavor in which academic institutions and industrial companies in different parts of the world collaborate. Therefore, this tutorial is relevant also to researchers and practitioners in countries that do not regulate AI technologies yet, in particular since we are experiencing more and more of such regulations recently.

Relevance to IR Community

We strongly believe that this tutorial is important to the entire IR and RS community. Since the major part of the audience has technical background, raising awareness of the ethical implications of their work and of the implications of recent regulations on research and development of IR and RS technologies is of utmost importance.

This tutorial is related to the following tutorials held earlier at similar venues:

- *Bias Issues and Solutions in Recommender System* by Jiawei Chen, Xiang Wang, Fuli Feng, and Xiangnan He at RecSys 2021[6]
- *Addressing Bias and Fairness in Search Systems* by Ruoyuan Gao and Chirag Shah at SIGIR 2021[7]
- *Towards Fair Federated Learning* by Zirui Zhou, Lingyang Chu, Yong Zhang, Lanjun Wang, Changxin Liu, and Jian Pei at KDD 2021[8]
- *Advances in Bias-aware Recommendation on the Web* by Ludovico Boratto and Mirko Marras at WSDM 2021[9]
- *Responsible AI in Industry: Practical Challenges and Lessons Learned* by Krishnaram Kenthapadi, Ben Packer, Mehrnoosh Sameki, and Nashlie Sephus at WWW 2021[10]
- *Bias Issues and Solutions in Recommender System* by Jiawei Chen, Xiang Wang, Fuli Feng, and Xiangnan He at WWW 2021[11]

While some of the topics we address in the tutorial at hand, in particular fairness and transparency, have been discussed in other tutorials already, our tutorial offers several unique characteristics. First, unlike others that commonly do not take an interdisciplinary perspective, we put a strong emphasis on providing such a perspective from different angles and stakeholders. Second, we connect our discussion to recent regulatory measures, in particular against the background of recent EU regulations. Third, since we have not held this tutorial before at other venues, we can contribute novel viewpoints and opinions, and different expertise on the subject, which we tailor to the SIGIR community. Despite the fact that this is a novel tutorial, we regularly cover the topics of ethics in information retrieval and recommendation systems in our lectures, interviews, and invited talks.

Format and Detailed Schedule

The tutorial is held as a 3-hour-tutorial plus additional breaks. The tutorial is organized into five parts: an introduction; three subsequent parts corresponding to the main themes addressed, i.e., fairness and non-discrimination, diversity, and transparency; and a discussion of open challenges. Throughout the three main parts, we discuss three perspectives: the system-centric perspective, the human-centric perspective, and the legal perspective, covering technical aspects, human needs, and legislators' points of view, respectively. More precisely, the tutorial covers the following aspects and is organized accordingly:

(1) **Introduction (15 minutes)**
Tutorial background, motivation, objectives, relevance to community, recent political and legal regulations

(2) **Fairness and non-discrimination (50 minutes)**
  (a) *Stakeholders:* We discuss the various stakeholders of retrieval and recommender systems, approaching the question for whom the system should be fair.
  (b) *Definition and quantification of bias and fairness:* We introduce the various kinds of bias and fairness concepts and definitions that are relevant for IR and RS research, along different axes (e.g., societal vs. statistical biases, model vs. presentation bias, provider vs. consumer fairness); we review the most common measures and metrics to quantify bias and fairness; we discuss their relation to political and legal regulations.
  (c) *Algorithms to mitigate biases and improve fairness:* We categorize the main strategies to mitigate harmful biases and improve fairness of retrieval and recommender systems, e.g., into pre-, in-, and post-processing techniques; we present concrete methods for each of these categories.
  (d) *Technical versus ethical and legal perspectives:* We discuss how the regulatory and legal frameworks align with the operationalization of fairness according to formal definitions often found in IR and RS papers.

(3) **Diversity (50 minutes)**
  (a) *Categories of diversity:* We introduce and discuss various kinds of diversity, i.e., personnel diversity in the research community and development teams, but also diversity in terms of the creators of content that can be retrieved or recommended.
  (b) *Diversity axes:* We elaborate on important groups or axes of diversity, including adults to children (age), from men to women to diverse genders, from western to non-western (culture), minority groups (e.g., indigenous people) and scientific disciplines.
  (c) *Diversity in the research community:* We present statistics of diversity aspects in the IR and RS communities, and ideas how to increase diversity.

[6] https://recsys.acm.org/recsys21/tutorials/#content-tab-1-5-tab
[7] https://sigir.org/sigir2021/tutorials
[8] https://kdd.org/kdd2021/tutorials
[9] https://www.wsdm-conference.org/2021/tutorials.php#2
[10] https://www2021.thewebconf.org/program/tutorials
[11] https://www2021.thewebconf.org/program/tutorials

(d) *Integrating diversity in evaluation:* We present strategies for considering diversity in the evaluation of IR and RS algorithms, in terms of adopted metrics, participants in user evaluations, and perspectives.

(4) **Transparency (50 minutes)**

(a) *Categories of transparency:* We introduce the major aspects of transparency, as they relate to building trust in IR and RS technology; we focus on explainability, traceability, and communication; we review and clarify the terminology.

(b) *Explainability and justification:* We discuss major strategies to achieve explainability of IR and RS technology, i.e., provide means to understand how the system works, targeting different stakeholders (e.g., developers vs. end users); we review approaches to provide justifications, i.e., mechanisms for the system to justify why a system outputs a certain (list of) documents or items.

(c) *Traceability and auditability:* We discuss strategies to keep track of the behavior of a system in a chronological way, in particular with the aim of facilitating auditing. We also point to recent works that discuss legal groundings and consequences of algorithmic auditing approaches, which is an underresearched topic to date [27].

(d) *Communication and logs:* We discuss the importance of documenting the development process, the resulting models, system capabilities, intended use, and limitations.

(5) **Open Challenges (15 minutes)**

(a) Understanding the discrepancy between (1) bias, fairness, and diversity metrics, (2) human perception of these aspects and factors influencing this perception, and (3) regulatory frameworks.

(b) Understanding the capabilities and limitations of existing solutions in terms of fairness, diversity, and transparency.

(c) Taking a multistakeholder perspective when developing solutions for fairness, diversity, and transparency in IR and RS technology.

(d) Improving the communication between the different stakeholders and between relevant research communities, including computer science, law, ethics, economy, sociology, psychology, in order to foster interdisciplinarity.

In order to *engage with both physical and virtual SIGIR attendees*, the tutorial includes time slots for audience interaction by means of surveys and opinion polls, brainstorming periods, and practical activities (e.g., applying some general concepts and requirements to specific IR use cases in terms of application domain, task, and user profile). We take advantage of collaborative tools such as Slido, Jamboard, or Padlet.

Supporting Material

The tutorial is supported by a GitHub repository containing an overview of the program with further details about the tutorial. The GitHub repository also contains the tutorial slides with references to all relevant works, software, and datasets. It can be accessed at https://github.com/socialcomplab/Retrieval-RecSys-AI-Ethics-Regulation-Tutorial-SIGIR22.

ACKNOWLEDGMENT

This work received financial support by the Austrian Science Fund (FWF): P33526 and DFH-23, by the State of Upper Austria and the Federal Ministry of Education, Science, and Research, through grant LIT-2020-9-SEE-113, and by the European Commission: HUMAINT project, Digital Economy Unit, Joint Research Centre.

REFERENCES

[1] Darius Afchar, Alessandro B Melchiorre, Markus Schedl, Romain Hennequin, Elena V Epure, and Manuel Moussallam. 2022. Explainability in Music Recommender Systems. *arXiv preprint arXiv:2201.10528* (2022).

[2] Eszter Bokányi and Anikó Hannák. 2020. Understanding inequalities in ride-hailing services through simulations. *Scientific reports* 10, 1 (2020), 1–11.

[3] Shea Brown, Jovana Davidovic, and Ali Hasan. 2021. The algorithm audit: Scoring the algorithms that score us. *Big Data & Society* 8, 1 (2021), 2053951720983865.

[4] Peter Brusilovsky, Marco de Gemmis, Alexander Felfernig, Elisabeth Lex, Pasquale Lops, Giovanni Semeraro, and Martijn C. Willemsen. 2021. Joint Workshop on Interfaces and Human Decision Making for Recommender Systems (IntRS'21). In *Fifteenth ACM Conference on Recommender Systems* (Amsterdam, Netherlands) *(RecSys '21)*. 783–786. https://doi.org/10.1145/3460231.3470927

[5] Robin Burke, Michael D Ekstrand, Nava Tintarev, and Julita Vassileva. 2021. Preface to the special issue on fair, accountable, and transparent recommender systems. *User Modeling and User-Adapted Interaction* 31, 3 (2021), 371–375.

[6] 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

[7] Yashar Deldjoo, Alejandro Bellogin, and Tommaso Di Noia. 2021. Explaining recommender systems fairness and accuracy through the lens of data characteristics. *Information Processing & Management* 58, 5 (2021), 102662.

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

[9] 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

[10] Michael D. Ekstrand, Anubrata Das, Robin Burke, and Fernando Diaz. 2021. Fairness and Discrimination in Information Access Systems. *CoRR* abs/2105.05779 (2021). arXiv:2105.05779 https://arxiv.org/abs/2105.05779

[11] European Parliament, Council, and Commission. 2012. Charter of Fundamental Rights of the European Union. *Official Journal of the European Union* C 326 (October 2012), 391–407. https://ec.europa.eu/info/aid-development-cooperation-fundamental-rights/your-rights-eu/eu-charter-fundamental-rights_en

[12] Ana Freire, Lorenzo Porcaro, and Emilia Gómez. 2021. Measuring Diversity of Artificial Intelligence Conferences. In *AAAI Workshop on Diversity in Artificial Intelligence (AIDBEI 2021)*.

[13] High-Level Expert Group on Artificial Intelligence. 2019. Ethics Guidelines for Trustworthy AI. (November 2019). https://doi.org/10.2759/346720

[14] 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

[15] Simone Kopeinik, Dominik Kowald, Ilire Hasani-Mavriqi, and Elisabeth Lex. 2017. Improving Collaborative Filtering Using a Cognitive Model of Human Category Learning. *The Journal of Web Science* 4, 2 (2017), 45–61.

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

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

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

[19] 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.

[20] 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

[21] 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.

[22] Elisabeth Lex, Dominik Kowald, and Markus Schedl. 2020. Modeling Popularity and Temporal Drift of Music Genre Preferences. *Transactions of the International Society for Music Information Retrieval* 3, 1 (2020).

[23] Elisabeth Lex, Dominik Kowald, Paul Seitlinger, Thi Ngoc Trang Tran, Alexander Felfernig, and Markus Schedl. 2021. Psychology-informed Recommender Systems. *Found. Trends Inf. Retr.* 15, 2 (2021), 134–242. https://doi.org/10.1561/1500000090

[24] 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

[25] 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. *Inf. Process. Manag.* 58, 5 (2021), 102666. https://doi.org/10.1016/j.ipm.2021.102666

[26] Alessandro B. Melchiorre, Eva Zangerle, and Markus Schedl. 2020. Personality Bias of Music Recommendation Algorithms. In *RecSys 2020: Fourteenth ACM Conference on Recommender Systems, Virtual Event, Brazil, September 22-26, 2020*, Rodrygo L. T. Santos, Leandro Balby Marinho, Elizabeth M. Daly, Li Chen, Kim Falk, Noam Koenigstein, and Edleno Silva de Moura (Eds.). ACM, 533–538. https://doi.org/10.1145/3383313.3412223

[27] Erwan Le Merrer, Ronan Pons, and Gilles Trédan. 2022. Algorithmic audits of algorithms, and the law. *arXiv preprint arXiv:2203.03711* (2022).

[28] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. Model cards for model reporting. In *Conference on fairness, accountability, and transparency*. 220–229.

[29] Peter Muellner, Dominik Kowald, and Elisabeth Lex. 2021. Robustness of Meta Matrix Factorization Against Strict Privacy Constraints. In *Advances in Information Retrieval*. Springer International Publishing, Cham, 107–119.

[30] Cataldo Musto, Marco de Gemmis, Pasquale Lops, and Giovanni Semeraro. 2021. Generating post hoc review-based natural language justifications for recommender systems. *User Modeling and User-Adapted Interaction* 31, 3 (2021), 629–673.

[31] Brandie Nonnecke and Camille Carlton. 2022. EU and US legislation seek to open up digital platform data. *Science* 375, 6581 (2022), 610–612. https://doi.org/10.1126/science.abl8537 arXiv:https://www.science.org/doi/pdf/10.1126/science.abl8537

[32] OECD. 2022. OECD Framework for the Classification of AI systems. 323 (2022). https://doi.org/https://doi.org/10.1787/cb6d9eca-en

[33] Inioluwa Deborah Raji, Andrew Smart, Rebecca N White, Margaret Mitchell, Timnit Gebru, Ben Hutchinson, Jamila Smith-Loud, Daniel Theron, and Parker Barnes. 2020. Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. 33–44.

[34] Markus Reiter-Haas, Emilia Parada-Cabaleiro, Markus Schedl, Elham Motamedi, Marko Tkalcic, and Elisabeth Lex. 2021. Predicting Music Relistening Behavior Using the ACT-R Framework. In *Fifteenth ACM Conference on Recommender Systems*. 702–707.

[35] 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

[36] Christian Sandvig, Kevin Hamilton, Karrie Karahalios, and Cedric Langbort. 2014. Auditing algorithms: Research methods for detecting discrimination on internet platforms. *Data and discrimination: converting critical concerns into productive inquiry* 22 (2014).

[37] Markus Schedl, Christine Bauer, Wolfgang Reisinger, Dominik Kowald, and Elisabeth Lex. 2021. Listener modeling and context-aware music recommendation based on country archetypes. *Frontiers in Artificial Intelligence* (2021), 108.

[38] Markus Schedl, David Hauger, Katayoun Farrahi, and Marko Tkalcic. 2015. On the Influence of User Characteristics on Music Recommendation Algorithms. In *Advances in Information Retrieval - 37th European Conference on IR Research, ECIR 2015, Vienna, Austria, March 29 - April 2, 2015. Proceedings (Lecture Notes in Computer Science, Vol. 9022)*, Allan Hanbury, Gabriella Kazai, Andreas Rauber, and Norbert Fuhr (Eds.). 339–345. https://doi.org/10.1007/978-3-319-16354-3_37

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

[40] Rashmi Sinha and Kirsten Swearingen. 2002. The role of transparency in recommender systems. In *CHI'02 extended abstracts on Human factors in computing systems*. 830–831.

[41] 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.

[42] Talia H Swartz, Ann-Gel S Palermo, Sandra K Masur, and Judith A Aberg. 2019. The Science and Value of Diversity: Closing the Gaps in Our Understanding of Inclusion and Diversity. *The Journal of Infectious Diseases* 220, Supplement 2 (08 2019), S33–S41. https://doi.org/10.1093/infdis/jiz174

[43] Nava Tintarev and Judith Masthoff. 2007. A survey of explanations in recommender systems. In *2007 IEEE 23rd international conference on data engineering workshop*. IEEE, 801–810.

[44] Sarah Myers West, Meredith Whittaker, and Kate Crawford. 2019. Discriminating systems. (2019).

[45] 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

[46] Yong Zheng and Juan Ruiz Toribio. 2021. The role of transparency in multi-stakeholder educational recommendations. *User Modeling and User-Adapted Interaction* 31, 3 (2021), 513–540.

[47] 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