

Fairness in Information Access

Conceptual Foundations and New Directions

Michael D. Ekstrand¹[0000-0003-2467-0108]

¹ Drexel University, Philadelphia PA 19104, USA
mdekstrand@drexel.edu

Abstract. As information access systems have profound impact through their mediation of users' experiences of information spaces, it is vital to ensure that their benefits, costs, and other impacts are *fair*: that people have equitable access to information, relevant information results, and opportunities, and are not systematically excluded, underserved, or harmed. Fairness is further a broad and complex topic with many divergent and sometimes competing definitions. This tutorial will provide an introduction to fairness in machine learning and information access, a broad survey of the landscape of fairness and fairness-related harms, and a review of key results and both established and emerging techniques for measuring and improving the fairness of information retrieval technologies.

Keywords: fairness, bias, evaluation.

1 Introduction

Information retrieval (IR) researchers have long been concerned with a range of different effects and views on system effectiveness and behavior, going beyond simple utility and aggregates thereof to consider topics such as novelty and diversity [1], political consequences and implications [2, 3], democratic principles [4], stereotypes [5, 6], and more. Common to several of these considerations is the need to understand how a search engine, recommender system, or other information access system interacts with, catalyzes, and responds to various *social factors*, such as bias, discrimination, and political processes; and to ensure that the system is aligned with the values and needs of humans, both individually and collectively [7].

Fairness is one family of these considerations, examining whether an information access system provides its benefits and opportunities in a manner that aligns with social ideas of what it means to be “fair”, such as distributing opportunity for user attention equitably among resources with comparable relevance to the user’s information need, or ensuring that no one group of people is systematically harmed or disadvantaged by the system’s operation or effects [8, 9].

In this tutorial, I present a roadmap of the broad space of fairness in information retrieval and related disciplines to help orient researchers new to fairness research, or new to IR research in general, to the various frameworks, research questions, and methods studied under the label of fairness in information access, and provide a review of

both established and newly-emerging techniques and methods for analyzing existing systems and developing new fair information access systems.

2 Intended Audience and Learning Outcomes

This tutorial is intended for researchers at all levels, with a particular attention to the needs of junior researchers (e.g. Ph.D. students), who would like a better understanding of the core principles, breadth of problems, and interesting new directions in fair information access. By the end of the tutorial, I hope that attendees are able to:

- Locate a fairness concern for information access in the broad landscape of fairness problems and questions, and describe its relationship to other problems.
- Identify data sets and experimental methods appropriate to study the concern.
- Reason about the different potential fairness problems that may emerge in a particular information retrieval application.
- Examine possible sources of unfairness in an application.

3 Outline

This tutorial is divided into two parts, based on the half-day tutorial format. The outline is based the monograph [8] and previous tutorials [10] co-developed with my collaborators, with the content significantly revised and updated based on newer directions and findings in the field.

3.1 Part 1: Definitions and Maps

The first part of the tutorial focuses on defining fairness and fairness-related problems, and mapping out the problem and solution spaces of fair information access.

1. Introduction — what is fairness?
 - a. Key terms, definitions, and relationships (fairness, bias, discrimination, etc.)
 - b. Social objectives and contestation [11]
 - c. Concreteness, abstractness, and “fairness-related harms”
2. Fairness in ML and AI
 - a. Key definitions, principles, and tradeoffs [12]
 - b. Sources of unfairness [13, 14]
 - c. Relationships between different fairness concepts
3. Mapping the Landscape of Fair Information Access [8]
 - a. Why information access is different from classification
 - b. Dimensions of fairness-related harms
 - c. Key considerations in defining a problem
4. Solution Spaces [9]
 - a. Data processing solutions
 - b. Algorithmic solutions

- c. Evaluation solutions
- d. Process and engineering solutions

3.2 Part 2: Topics and Findings

The second half of the tutorial surveys the literature on various aspects of fairness, describing key findings and methods in the context of the landscape of Part 1.

5. Consumer-side Fairness [9, 15]
 - a. Definitions and core problems
 - b. Subtractibility and zero- vs. positive-sum metrics
 - c. Consequences of consumer-unfair systems
 - d. Stereotypes and representational harms [e.g., 6]
6. Provider-Side Fairness
 - a. Definitions and metrics
 - b. Methods for improving provider-side fairness
 - c. Feedback loops and ecosystem effects
7. Other Stakeholders and Considerations
 - a. Subject-side fairness
 - b. Multi-stakeholder fairness
8. Resources for Research and Practice
9. Conclusion

4 Post-Tutorial Resources

Tutorial slides and a complete bibliography, along with a recording if one is provided by the conference, will be available during and after the tutorial from my web site.¹

Acknowledgments. This tutorial is based on work supported by the National Science Foundation under Grant Nos. 17-51278 and 24-15042. I gratefully acknowledge my many collaborators who have contributed to the work I present and sharpened my thinking on this topic, and in particular Robin Burke and Fernando Diaz, who co-authored our earlier series of tutorials, and Anubrata Das, who joined us for our monograph on fairness in information access.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article

References

1. Clarke, C.L.A., Kolla, M., Cormack, G.V., Vechtomova, O., Ashkan, A., Büttcher, S., MacKinnon, I.: Novelty and diversity in information retrieval evaluation. In: Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in

¹ <https://md.ekstrandom.net/talks/2025/ecir-fairness-tutorial>

- Information Retrieval (SIGIR '08). pp. 659–666. ACM (2008). <https://doi.org/10.1145/1390334.1390446>.
2. Belkin, N.J., Robertson, S.E.: Some ethical and political implications of theoretical research in information science. In: Proceedings of the ASIS Annual Meeting (1976).
 3. Introna, L.D., Nissenbaum, H.: Shaping the Web: why the politics of search engines matters. *The Information Society*. 16, 169–185 (2000). <https://doi.org/10.1080/01972240050133634>.
 4. Vrijenhoek, S., Kaya, M., Metoui, N., Möller, J., Odijk, D., Helberger, N.: Recommenders with a mission: assessing diversity in news recommendations. In: Proceedings of the 2021 Conference on Human Information Interaction and Retrieval (CHIIR '21). pp. 173–183. ACM (2021). <https://doi.org/10.1145/3406522.3446019>.
 5. Bolukbasi, T., Chang, K.-W., Zou, J., Saligrama, V., Kalai, A.: Quantifying and Reducing Stereotypes in Word Embeddings, <http://arxiv.org/abs/1606.06121>, (2016).
 6. Raj, A., Ekstrand, M.D.: Fire dragon and unicorn princess: gender stereotypes and children’s products in search engine responses. In: Proceedings of the 2022 SIGIR Workshop on eCommerce (2022).
 7. Stray, J., Halevy, A., Assar, P., Hadfield-Menell, D., Boutilier, C., Ashar, A., Bakalar, C., Beattie, L., Ekstrand, M., Leibowicz, C., Moon Sehat, C., Johansen, S., Kerlin, L., Vickrey, D., Singh, S., Vrijenhoek, S., Zhang, A., Andrus, M., Helberger, N., Proutskova, P., Mitra, T., Vasani, N.: Building human values into recommender systems: an interdisciplinary synthesis. *ACM Trans. Recomm. Syst.* 2, 20:1-57 (2024). <https://doi.org/10.1145/3632297>.
 8. Ekstrand, M.D., Das, A., Burke, R., Diaz, F.: Fairness in information access systems. *Foundations and Trends® in Information Retrieval*. 16, 1–177 (2022). <https://doi.org/10.1561/15000000079>.
 9. Ekstrand, M.D., Beattie, L., Pera, M.S., Cramer, H.: Not just algorithms: strategically addressing consumer impacts in information retrieval. In: Proceedings of the 46th European Conference on Information Retrieval (ECIR '24). Springer (2024). https://doi.org/10.1007/978-3-031-56066-8_25.
 10. Ekstrand, M.D., Burke, R., Diaz, F.: Fairness and Discrimination in Retrieval and Recommendation. In: Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'19). pp. 1403–1404. ACM (2019). <https://doi.org/10.1145/3331184.3331380>.
 11. Selbst, A.D., Boyd, D., Friedler, S.A., Venkatasubramanian, S., Vertesi, J.: Fairness and abstraction in sociotechnical systems. In: Proceedings of the 2nd ACM Conference on Fairness, Accountability, and Transparency (FAT* '19). pp. 59–68. ACM (2019). <https://doi.org/10.1145/3287560.3287598>.
 12. Mitchell, S., Potash, E., Barocas, S., D’Amour, A., Lum, K.: Algorithmic Fairness: Choices, Assumptions, and Definitions. *Annual Review of Statistics and Its Application*. 8, (2020). <https://doi.org/10.1146/annurev-statistics-042720-125902>.
 13. Chen, I., Johansson, F.D., Sontag, D.: Why is my classifier discriminatory? In: Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R. (eds.) *Advances in Neural Information Processing Systems* 31. pp. 3539–3550. Curran Associates, Inc. (2018).
 14. Barocas, S., Selbst, A.D.: Big data’s disparate impact. *California Law Review*. 104, 671 (2016). <https://doi.org/10.15779/Z38BG31>.
 15. Boratto, L., Fenu, G., Marras, M., Medda, G.: Consumer fairness in recommender systems: contextualizing definitions and mitigations. In: *Advances in information retrieval: ECIR 2022*. pp. 552–566. Springer (2022). https://doi.org/10.1007/978-3-030-99736-6_37.