NORMalize 2024: The Second workshop on normative design and evaluation of recommender systems

Starke, A.D.; Vrijenhoek, S.; Michiels, L.; Kruse, J.; Tintarev, N.

DOI
10.1145/3640457.3687103

Publication date
2024

Document Version
Final published version

Published in
RecSys '24

NORMalize 2024: The Second Workshop on Normative Design and Evaluation of Recommender Systems

Alain D. Starke
a.d.starke@uva.nl
ASCoR, University of Amsterdam
Amsterdam, The Netherlands
MediaFutures, University of Bergen
Bergen, Vestland, Norway

Sanne Vrijenhoek
s.vrijenhoek@uva.nl
University of Amsterdam
Amsterdam, The Netherlands
AI, Media & Democracy Lab
Amsterdam, The Netherlands

Lien Michiels
lien.michiels@uantwerpen.be
University of Antwerp
Antwerp, Belgium
imec-SMIT, Vrije Universiteit Brussel
Brussels, Belgium

Johannes Kruse
johannes.kruse@jppol.dk
JP/Politikens Media Group
Copenhagen, Denmark
Technical University of Denmark
Kongens Lyngby, Denmark

Nava Tintarev
n.tintarev@maastrichtuniversity.nl
Maastricht University
Maastricht, The Netherlands

ABSTRACT

Recommender systems are among the most widely used applications of artificial intelligence. Their use can have far-reaching consequences for users, stakeholders, and society at large. In this second edition of the NORMalize workshop, we once again seek to advance the research agenda of *normative thinking*, considering the norms and values that underpin recommender systems, as well as to introduce the concept to a broader audience. We aim to bring together a growing community of researchers and practitioners across disciplines who want to think about the norms and values that should be considered in the design and evaluation of recommender systems, and to further educate them on how to reflect on, prioritise, and operationalise such norms and values. NORMalize 2024 is a half-day workshop consisting of a combination of paper presentations and an interactive session, building upon its successful full-day run last year at RecSys'23.

CCS CONCEPTS

• **Information systems → Recommender systems**; • **Social and professional topics → Systems analysis and design**.

KEYWORDS

normative thinking, normative design, norms, values, value-sensitive design

ACM Reference Format:
Alain D. Starke, Sanne Vrijenhoek, Lien Michiels, Johannes Kruse, and Nava Tintarev. 2024. NORMalize 2024: The Second Workshop on Normative Design and Evaluation of Recommender Systems. In *18th ACM Conference on Recommender Systems (RecSys '24), October 14–18, 2024, Bari, Italy.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3640457.3687103

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

ACM ISBN 979-8-4007-0505-2/24/10
https://doi.org/10.1145/3640457.3687103

1 INTRODUCTION

Building upon last year's edition, NORMalize 2024 will once again promote normative thinking and decision-making about the norms and values that underpin the design and evaluation of recommender systems with the ACM RecSys conference's predominantly technical audience. NORMalize thus takes a distinct position in the extensive landscape of workshops at the ACM RecSys conference.

NORMalize builds upon extensive knowledge regarding *normative* thinking and decision-making [1, 3, 21], which is a pillar of humanities and social science research, but only recently has been attracting interest from the exact sciences, specifically, regarding applications of machine learning and artificial intelligence.

Normative thinking requires researchers and practitioners to go beyond the current state of the system (output) and instead reflect on how or what the system *should be*. This involves identifying a system's relevant values, determining how these values would be expressed in a system's recommended content, examining how different values may conflict, and justifying how certain values in such cases should be prioritized over others [17]. For example, company values or societal norms might steer a recommender's algorithm to mainly promote pro-environmental options over other options (cf. [18, 20]).

Importantly, normative thinking does not lead to general design standards for recommender systems. Instead, it provides the tools to develop a normative claim *for a system specifically*, given its operating environment and goals. News recommender systems are a notable example, for they essentially 'gatekeep' which articles and whose voices are presented to the general public [9]. News organizations (un)consciously set normative recommendation goals based on their perceived role in society. For example, 'informing' would call for timely coverage of important events, 'educating' for an increased exposure to views different from a reader's own, and 'entertaining' to cater to user preferences only. Such recommender goals and values, all equally valid, can be at odds with each other or even mutually exclusive [11].

Similar decisions need to be made for other domains as well. For example, developers of food recommenders could promote the

consumption of nutritious food due to health pressures in society [8, 19]. They should aim to balance this with personal preferences and economic incentives, as not all users have (financial) access to items satisfying both criteria, such as locally produced blueberries. Other domains, such as music recommendation [10], job matching [23], dating apps [2] and tourism [4] each come with their own normative challenges.

Recently, we have seen an increase in interest in normative thinking with the RecSys community [7, 12–15, 18, 22, 24]. However, leveraging the decades worth of 'normative knowledge' built up within the humanities and social sciences is difficult, due to there being few opportunities for interdisciplinary collaboration.

Key Outcomes

Similar to previous editions, NORMalize 2024 thus seeks to foster a space for interdisciplinary discussion between the humanities, social sciences, and exact sciences about the norms and values underlying the design and evaluation of recommender systems. Also like previous editions, NORMalize 2024 welcomes both the 'norm curious' and the 'norm active', by providing opportunities for both knowledge sharing and discussion. In doing so, we aim to stimulate further interdisciplinary collaboration, which will ultimately lead to better normative design and evaluation of recommender systems.

2 ORGANISER BIOGRAPHIES

The organisers of this workshop cover different domains, backgrounds, and experience levels. Their combined networks and experiences give this workshop the potential to be exceptionally interdisciplinary.

Alain D. Starke is an assistant professor in persuasive communication for a digital society, at the University of Amsterdam, Netherlands. He is also an adjunct associate professor in recommender systems at the SFI MediaFutures research centre for responsible media technology, which is part of the University of Bergen, Norway. His research aims to develop recommender systems that can support preference shifts and behavioural change in domains of self-actualisation, such as energy conservation, healthy eating, and news recommendation.

Sanne Vrijenhoek is a PhD Candidate at the University of Amsterdam's Institute of Information Law with a background in Artificial Intelligence. She works on an interdisciplinary project assessing diversity in news recommendations. An important part of this project is translating normative notions of diversity into concrete concepts that can be used to inform recommender system design. Her work was awarded Best Paper Runner Up at RecSys'22 [24].

Lien Michiels is a PhD Candidate in the Adrem Data Lab at the University of Antwerp, Belgium and a researcher at imec-SMIT, Vrije Universiteit Brussel. She is the lead researcher on the FWO SBO funded 'Serendipity Engine' project for the Adrem Data Lab. As part of this project, she applies normative design principles to urban and news recommender systems leading to more diverse and serendipitous experiences for users. Previously, she combined her PhD research with her work as a Machine Learning Engineer at Froomle.

Johannes Kruse is an industrial PhD Candidate at the Technical University of Denmark's Department of Applied Mathematics and Computer Science in collaboration with the Danish news publishers Ekstra Bladet and JP/Politikens Media Group. He is in charge of developing and maintaining the core recommendation systems at EkstraBladet.dk, which serve millions of users. He focuses on creating algorithms that provide personalised recommendations while balancing relevance and diversity.

Nava Tintarev is a full professor in explainable AI in the Department of Advanced Computing Sciences at Maastricht University, Netherlands. Her research looks at how to improve transparency in, and decision support for, recommender systems. She is a Co-Investigator in the ROBUST consortium carrying out long-term (10-year) research into trustworthy artificial intelligence. She is also a co-lab director of the TAIM lab, working on trustworthy media, in collaboration with UvA and RTL. Her recent work on a.o., diversification of news and social media items has received four best paper awards in the last four years [5, 6, 16, 25].

3 WORKSHOP FORMAT

NORMalize 2024 is a half-day workshop consisting of two parts. In the first part, authors of accepted contributions will have the opportunity to present and discuss their work. In the second part, we will facilitate an interactive session where participants will practice normative reasoning and constructive disagreement through so-called 'disagreenotes'. Disagreenotes consists of normative statements that we expect participants will have different perspectives on, e.g., "the dangers of personalization in the domain of recruiting are such that we should avoid using it, if possible". The full workshop program can be found on the workshop website: https://sites.google.com/view/normalizeworkshop.

4 CONTRIBUTIONS

NORMalize 2024 called for contributions in the form of full papers (12 pages), short papers (6 pages), and extended abstracts (3 pages). The topics of interest included, but were not limited to: (1) *Defining normativity for different stakeholders and in various recommender domains*, (2) *Frameworks for formalizing metrics and generalizing across domains*, (3) *Value-sensitive design*, (4) *Survey papers related to normativity*, (5) *Algorithmic improvements for normative values*, (6) *Using datasets and data representation for normativity*, (7) *Case studies from practitioners on solving problems related to normativity*, (8) *Balancing multiple objectives and stakeholder optimization with different norms and values*, (9) *Empirical analyses measuring recommender systems behavior using simulations, offline, or online studies*. All accepted contributions will be presented at the workshop and published as CEUR-WS Proceedings.

The reviewing process for NORMalize 2024 was mutually anonymous (double-blind). Each paper was reviewed by three members of our program committee of at least two different backgrounds, reflecting the interdisciplinary character of NORMalize 2024.

5 PROGRAM COMMITTEE

We are grateful to our program committee, made up of scholars and practitioners from a wide range of backgrounds and expertises, for their invaluable contributions to the success of NORMalize 2024.

- **Exact Sciences:** Christine Bauer (Paris Lodron University of Salzburg), Dietmar Jannach (University of Klagenfurt), Joel Pinho Lucas (Globo), Julia Neidhardt (Vienna University of Technology), Laura Jansen (Wageningen University), Leticia Freire Figueiredo (Universidade Federal Fluminense), Lorenzo Porcaro (European Centre for Algorithmic Transparency), Lucien Heitz (University of Zurich), Manel Slokom (Centrum Wiskunde & Informatica), Myrthe Reuver (VU Amsterdam), Robin Verachtert (DPG Media), Savvina Daniil (Centrum Wiskunde & Informatica)
- **Social Sciences:** Damian Trilling (University of Amsterdam), Kasper Lindskow (JP/Politikens Media Group), Kimon Kieslich (University of Amsterdam), Laura Modre (University of Vienna), Nicolas Mattis (VU Amsterdam), Rupert Kiddle (University of Amsterdam), Sophie Morosoli (University of Amsterdam), Zilin Lin (University of Amsterdam)
- **Humanities & Law:** Ljubiša Metikoš (University of Amsterdam), Louisa Bartolo (Flint Global), Marijn Sax (University of Amsterdam), Natali Helberger (University of Amsterdam), Theresa Seipp (University of Amsterdam)

ACKNOWLEDGMENTS

We would like to thank the participants and authors of accepted contributions for their valuable inputs to the workshop. Finally, we would like to thank our employers and funding bodies. Sanne Vrijenhoek's contribution to this research is supported by the AI, Media and Democracy Lab. Lien Michiels' contribution to this research was supported by the Research Foundation Flanders (FWO) under grant number S006323N and the Flanders AI Research Program. Johannes Kruse's contribution to this research is supported by the Innovation Foundation Denmark under grant number 1044-00058B and Platform Intelligence in News under project number 0175-00014B. Alain Starke's contribution was in part supported by the Research Council of Norway with funding to MediaFutures: Research Centre for Responsible Media Technology and Innovation, through the Centre for Research-based Innovation scheme, project number 309339. Nava Tintarev's contribution is supported by the project ROBUST: Trustworthy AI-based Systems for Sustainable Growth with project number KICH3.LTP.20.006, which is (partly) financed by the Dutch Research Council (NWO), RTL, and the Dutch Ministry of Economic Affairs and Climate Policy (EZK) under the program LTP KIC 2020-2023. *All content represents the opinion of the authors, which is not necessarily shared or endorsed by their respective employers and/or sponsors.*

REFERENCES

[1] Steve Buckler. 2010. Normative theory. *Theory and methods in political science* 3 (2010), 156–180.
[2] Paul Byron and Kath Albury. 2018. 'There are literally no rules when it comes to these things': Ethical practice and the use of dating/hook-up apps. *Digital intimate publics and social media* (2018), 213–229.
[3] Theresia Anita Christiani. 2016. Normative and empirical research methods: Their usefulness and relevance in the study of law as an object. *Procedia-Social and Behavioral Sciences* 219 (2016), 201–207.
[4] Manoj Reddy Dareddy. 2016. Challenges in Recommender Systems for Tourism.. In *RecTour@ RecSys*. 59–61.
[5] Tim Draws, Oana Inel, Nava Tintarev, Christian Baden, and Benjamin Timmermans. 2022. Comprehensive viewpoint representations for a deeper understanding of user interactions with debated topics. In *Proceedings of the 2022 Conference on Human Information Interaction and Retrieval*. 135–145.
[6] Tim Draws, Alisa Rieger, Oana Inel, Ujwal Gadiraju, and Nava Tintarev. 2021. A checklist to combat cognitive biases in crowdsourcing. In *Proceedings of the AAAI conference on human computation and crowdsourcing*, Vol. 9. 48–59.
[7] Michael D Ekstrand, Mucun Tian, Mohammed R Imran Kazi, Hoda Mehrpouyan, and Daniel Kluver. 2018. Exploring author gender in book rating and recommendation. In *Proceedings of the 12th ACM conference on recommender systems*. 242–250.
[8] Hanna Hauptmann, Alan Said, and Christoph Trattner. 2022. Research directions in recommender systems for health and well-being: A Preface to the Special Issue. *User Modeling and User-Adapted Interaction* (2022), 1–6.
[9] Natali Helberger. 2019. On the democratic role of news recommenders. *Digital Journalism* 7, 8 (2019), 993–1012.
[10] Andre Holzapfel, Bob Sturm, and Mark Coeckelbergh. 2018. Ethical dimensions of music information retrieval technology. *Transactions of the International Society for Music Information Retrieval* 1, 1 (2018), 44–55.
[11] Jon Kleinberg, Sendhil Mullainathan, and Manish Raghavan. 2016. Inherent trade-offs in the fair determination of risk scores. *arXiv preprint arXiv:1609.05807* (2016).
[12] Bart P Knijnenburg, Saadhika Sivakumar, and Daricia Wilkinson. 2016. Recommender systems for self-actualization. In *Proceedings of the 10th acm conference on recommender systems*. 11–14.
[13] 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*. 2243–2251.
[14] Lien Michiels, Jorre Vannieuwenhuyze, Jens Leysen, Robin Verachtert, Annelien Smets, and Bart Goethals. 2023. How Should We Measure Filter Bubbles? A Regression Model and Evidence for Online News. In *Proceedings of the 17th ACM Conference on Recommender Systems* (Singapore, Singapore) *(RecSys '23)*. Association for Computing Machinery, New York, NY, USA, 640–651. https://doi.org/10.1145/3604915.3608805
[15] Erasmo Purificato, Ludovico Boratto, and Ernesto William De Luca. 2022. Do Graph Neural Networks Build Fair User Models? Assessing Disparate Impact and Mistreatment in Behavioural User Profiling. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*. 4399–4403.
[16] Alisa Rieger, Tim Draws, Mariët Theune, and Nava Tintarev. 2021. This item might reinforce your opinion: Obfuscation and labeling of search results to mitigate confirmation bias. In *Proceedings of the 32nd ACM conference on hypertext and social media*. 189–199.
[17] Bernd Carsten Stahl. 2012. Morality, ethics, and reflection: a categorization of normative IS research. *Journal of the association for information systems* 13, 8 (2012), 1.
[18] Alain Starke, Martijn Willemsen, and Chris Snijders. 2021. Promoting energy-efficient behavior by depicting social norms in a recommender interface. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 11, 3-4 (2021), 1–32.
[19] Alain D Starke, Cataldo Musto, Amon Rapp, Giovanni Semeraro, and Christoph Trattner. 2024. "Tell Me Why": using natural language justifications in a recipe recommender system to support healthier food choices. *User Modeling and User-Adapted Interaction* 34, 2 (2024), 407–440.
[20] Alain D Starke and Martijn C Willemsen. 2024. Psychologically Informed Design of Energy Recommender Systems: Are Nudges Still Effective in Tailored Choice Environments? In *A Human-Centered Perspective of Intelligent Personalized Environments and Systems*. Springer, 221–259.
[21] Judith Jarvis Thomson. 2010. Normativity.
[22] Christoph Trattner and David Elsweiler. 2019. What online data say about eating habits. *Nature Sustainability* 2, 7 (2019), 545–546.
[23] Elmira van den Broek, Anastasia Sergeeva, and Marleen Huysman. 2019. Hiring algorithms: An ethnography of fairness in practice. (2019).
[24] Sanne Vrijenhoek, Gabriel Bénédict, Mateo Gutierrez Granada, Daan Odijk, and Maarten De Rijke. 2022. RADio – Rank-Aware Divergence Metrics to Measure Normative Diversity in News Recommendations. In *Proceedings of the 16th ACM Conference on Recommender Systems* (Seattle, WA, USA) *(RecSys '22)*. Association for Computing Machinery, New York, NY, USA, 208–219. https://doi.org/10.1145/3523227.3546780
[25] Mireia Yurrita, Tim Draws, Agathe Balayn, Dave Murray-Rust, Nava Tintarev, and Alessandro Bozzon. 2023. Disentangling Fairness Perceptions in Algorithmic Decision-Making: the Effects of Explanations, Human Oversight, and Contestability. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–21.