



FATE/MM 20: 2nd International Workshop on Fairness, Accountability, Transparency and Ethics in MultiMedia

Xavier Alameda-Pineda
Inria, Perception Team
Grenoble, France
xavier.alameda-pineda@inria.fr

Miriam Redi
Wikimedia Foundation
San Francisco, California
mredi@wikimedia.org

Jahna Otterbacher
Open University of Cyprus
Latsia, Cyprus
jahna.otterbacher@ouc.ac.cy

Nicu Sebe
University of Trento
Trento, Italy
niculae.sebe@unitn.it

Shih-Fu Chang
Columbia University
New York, New York
sfchang@cs.columbia.edu

ABSTRACT

The series of FAT/FAccT events aim at bringing together researchers and practitioners interested in fairness, accountability, transparency and ethics of computational methods. The FATE/MM workshop focuses on addressing these issues in the Multimedia field. Multimedia computing technologies operate today at an unprecedented scale, with a growing community of scientists interested in multimedia models, tools and applications. Such continued growth has great implications not only for the scientific community, but also for the society as a whole. Typical risks of large-scale computational models include model bias and algorithmic discrimination. These risks become particularly prominent in the multimedia field, which historically has been focusing on user-centered technologies. To ensure a healthy and constructive development of the best multimedia technologies, this workshop offers a space to discuss how to develop ethical, fair, unbiased, representative, and transparent multimedia models, bringing together researchers from different areas to present computational solutions to these issues.

CCS CONCEPTS

• **Information systems**; • **Security and privacy**; • **Social and professional topics**; • **Human-centered computing**;

ACM Reference Format:

Xavier Alameda-Pineda, Miriam Redi, Jahna Otterbacher, Nicu Sebe, and Shih-Fu Chang. 2020. FATE/MM 20: 2nd International Workshop on Fairness, Accountability, Transparency and Ethics in MultiMedia. In *Proceedings of the 28th ACM International Conference on Multimedia (MM '20)*, October 12–16, 2020, Seattle, WA, USA. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3394171.3421896>

1 INTRODUCTION

The computational inclusiveness and transparency of automatic information processing methods is a research topic that exhibited

growing interest in the past years. In the era of digitized decision-making software where the push for artificial intelligence happens worldwide and at different strata of the socio-economic tissue, the consequences of biased, unexplainable and opaque methods for multimedia analysis and content retrieval, can be dramatic. Information processing is one of the fundamental pillars of multimedia: whether data is processed for content, experience or systems applications, the automatic processing of information is used in every corner of our community. In this context, the multimedia community must put together the necessary efforts in applying its expertise and know-how and investigate how to transform the current technical tools and methodologies so as to derive computational models that are ethical, transparent and inclusive.

2 TOPICS

The workshop aims to foster research around a timely and crucial topic for the present digitized society: the fairness, accountability, transparency and ethics of multimedia algorithms. The workshop has a strong scientific link with the FAT/ML workshop and the ACM FAccT conference. Differently from FAT/ML, which is anchored in machine learning, the FATE/MM workshop addresses fairness, accountability and transparency in the core of the multimedia community. We invited submissions covering any topic closely related to the multimedia community and falling in the following categories:

Models. : Techniques and models for fairness-aware multimedia modeling, retrieval, and recommendation; Interpretable and explainable models in multimedia; Models and frameworks for conducting FAT audits of multimedia systems; Models for addressing inclusion and exclusion in multimedia.

Algorithm evaluation. Qualitative, quantitative, and experimental studies on subjective perceptions of algorithmic bias and unfairness; Experimental results of FAT audits of multimedia systems; Objective metrics for measuring unfairness and bias; Generation of human-readable explanations for multimedia models and algorithmic outputs; Metrics for measuring inclusiveness in multimedia systems.

Data collection and curation. Defining, measuring and mitigating problematic biases in multimedia datasets; Improvement of data

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

MM '20, October 12–16, 2020, Seattle, WA, USA
© 2020 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-7988-5/20/10.
<https://doi.org/10.1145/3394171.3421896>

collection processes to be more fair, diverse, and inclusive; Data collection regarding potential unfairness in systems.

Applications. Research on fair and transparent multimedia tools and applications; Ethical design and/or usage of multimedia tools and applications.

3 PROGRAM

The program will follow the next guidelines:

- Invited talk: Prof. Vivek K. Singh
- Oral presentation #1
- Oral presentation #2
- Invited talk: Prof. Katharina Reinecke
- Coffee break
- Invited talk: Dr. Cristian Canton
- Oral presentation #3
- Oral presentation #4

There are four papers accepted at the workshop with the following titles: Gender Slopes: Counterfactual Fairness for Computer Vision Models by Attribute Manipulation; Not Judging a User by Their Cover: Understanding Harm in Multi-Modal Processing within Social Media Research; Fighting Filterbubbles with Adversarial Training; Balancing Fairness and Accuracy in Sentiment Detection using Multiple Black Box Models.

The first paper proposes an encoder-decoder network for image attribute manipulation which synthesizes facial images by varying only the dimensions of gender and race, leaving other traits intact. Such synthesized images are then used measure counterfactual fairness of commercial computer vision classifiers by examining. To do so, the authors compute the degree to which these classifiers are affected by gender and racial cues controlled in the images. For example, faces having feminine traits may elicit higher scores for the concept of nurse and lower scores for STEM-related concepts.

The second paper deals with fairness in the social media ecosystems, specifically focusing on the need for the social media community to understand multimedia processing and its unique ethical considerations. The authors design a set of crowdsourcing experiments for race, gender, and age annotation of Twitter users. Annotators are shown Twitter users' profiles under different modalities (text, image, both). Statistical differences are identified in performances of MTurk annotators when different modalities of information are provided, and the consequences of those biases are discussed.

The third paper deals with the role of recommender systems in the generation of filter bubbles, with a specific focus on news articles. The authors identify the root of the issue into the vectorial feature representation of the news text, and propose a new training scheme based on adversarial machine learning to tackle this issue. Preliminary experiments show that the features extracted with this method are significantly offer the possibility to reduce the risk of manifestation of new filter bubbles.

The last paper is about sentiment detection in multimedia computing. The main idea of the paper is to combine different modalities not only to improve accuracy, but also to increase fairness of the resulting sentiment classification, with the aim of building fair yet accurate sentiment detectors for multiple applications. The authors audit multiple commercial sentiment detection APIs for the gender bias in two-actor news headlines settings and report on the level

of bias observed. They then propose a "Flexible Fair Regression" approach, to ensure satisfactory accuracy and fairness by jointly learning from multiple black-box models.

4 ORGANIZERS

Our organization committee is very diverse in terms of expertise, topics, geographic location and seniority.

Dr. Xavier Alameda-Pineda is a Research Scientist at Inria Grenoble. He received his PhD in mathematics and Computer Science from Université Joseph Fourier in 2013. His main research interests are machine learning for computer vision and multi-modal fusion with applications to robotics and human behavior understanding. He co-organized several workshops at ACM Multimedia, ECCV and CVPR over the past few years.

Dr. Miriam Redi is a Senior Research Scientist at the Wikimedia Foundation and Visiting Research Fellow at King's College London. Formerly, she worked as a Research Scientist at Yahoo Labs in Barcelona and Nokia Bell Labs in Cambridge. She received her PhD from EURECOM, Sophia Antipolis. She conducts research in social multimedia computing, working on fair, interpretable, multimodal machine learning solutions to improve knowledge equity. She is the recipient of 2 paper awards, including ICWSM 2015 and ACM ICMR 2016.

Prof. Jahna Otterbacher (Ph.D., University of Michigan at Ann Arbor) is currently Assistant Professor at the Open University of Cyprus, Faculty of Pure and Applied Sciences, where she coordinates the Cyprus Center for Algorithmic Transparency (CyCAT). Jahna holds a concurrent appointment as team leader of the Transparency in Algorithms Group at RISE (Research centre on Interactive media, Smart systems and Emerging technologies), a new center of excellence and innovation in Nicosia, Cyprus.

Prof. Nicu Sebe is Full Professor with the University of Trento, Italy, leading the research in the areas of multimedia information retrieval and human behavior understanding. He was the General Co-Chair of the IEEE FG Conference 2008 and ACM Multimedia 2013, and the Program Chair of the International Conference on Image and Video Retrieval in 2007 and 2010, ACM Multimedia 2007 and 2011.

Prof. Shih-Fu Chang is the Sr. Executive Vice Dean and the Richard Dicker Professor of Columbia Engineering. His research is focused on multimedia information retrieval, computer vision, machine learning, and signal processing, with the goal to turn unstructured multimedia data into searchable information. For his long-term pioneering contributions, he has been awarded the IEEE Signal Processing Society Technical Achievement Award, ACM Multimedia SIG Technical Achievement Award, Honorary Doctorate from the University of Amsterdam, the IEEE Kiyo Tomiyasu Award, and IBM Faculty Award. He is a Fellow of the American Association for the Advancement of Science (AAAS) and IEEE.