Integrity 2023: Integrity in Social Networks and Media

Lluís Garcia-Pueyo
Meta
Menlo Park, USA
lgp@fb.com

Panayiotis Tsaparas
University of Ioannina
Ioannina, Greece
tsap@cs.uoi.gr

Prathyusha Senthil Kumar
Meta
Menlo Park, USA
prathyushas@fb.com

Timos Sellis
Archimedes / Athena Research Center
Athens, Greece
timos@athenarc.gr

Paolo Papotti
EURECOM
Biot, France
paolo.papotti@eurecom.fr

Sibel Adalı
Rensselaer Polytechnic Institute
New York, USA
sibel@cs.rpi.edu

Giuseppe Manco
ICAR-CNR
Rende, Italy
giuseppe.manco@icar.cnr.it

Tudor Trufinescu
Meta
Bellevue, USA
tudort@fb.com

Gireeja Ranade
UC Berkeley
Berkeley, USA
ranade@eecs.berkeley.edu

James Verbus
LinkedIn
Sunnyvale, USA
jverbus@linkedin.com

Mehmet N Tek
Google
Redwood City, USA
mtek@google.com

Anthony McCosker
Swinburne Social Innovation
Research Institute
Melbourne, Australia
amccosker@swin.edu.au

ABSTRACT

Integrity 2023 is the fourth edition of the successful Workshop on Integrity in Social Networks and Media, held in conjunction with the ACM Conference on Web Search and Data Mining (WSDM) in the past three years. The goal of the workshop is to bring together researchers and practitioners to discuss content and interaction integrity challenges in social networks and social media platforms. The event consists of a combination of invited talks by reputed members of the Integrity community from both academia and industry and peer-reviewed contributed talks and posters solicited via an open call-for-papers.

1 WORKSHOP DESCRIPTION

In the past decade, social networks and social media, such as Facebook and Twitter, have become the default channels of communication and information. The popularity of these online portals has exposed a collection of integrity issues; cases where the content produced and exchanged compromises the quality, operation, and eventually the integrity of the platform. Examples include misinformation, low quality and abusive content and behaviors, and polarization and opinion extremism. There is an urgent need to detect and mitigate the effects of these integrity issues in a timely, efficient, and unbiased manner.

The Integrity 2023 workshop aims to bring together top researchers and practitioners from academia and industry to engage in a discussion about algorithmic and systems aspects of integrity challenges, leading to exchange of knowledge and cross-cutting collaborations. The workshop topics include, but are not limited to:

- **Low quality, borderline, and offensive content and behaviors:** Methods for detecting and mitigating low quality and offensive content and behaviors, such as clickbait, fake engagement, nudity and violence, bullying, and hate speech.
- **Personalized treatment of low quality content:** Identification, measurement, and reduction of bad experiences.
- **COVID-19 on social media:** Authoritative health information; Covid misinformation; Vaccine hesitancy; Anti-vax movements.
- **Misinformation:** Detecting and combating misinformation; Prevalence and virality of misinformation; Misinformation sources and origins; Source and content credibility; Inoculation strategies; Deep and shallow fakes.
- **Polarization:** Models and metrics for polarization; Echo chambers and filter bubbles; Opinion Extremism and radicalization; Algorithms for mitigating polarization.
- **Fairness in Integrity:** Fairness in the detection and mitigation of integrity issues with respect to sensitive attributes such as gender, race, sexual orientation, and political affiliation.

There is a strong interest in the community in integrity, with several workshops and conferences on related topics [1–3, 7]. The Integrity 2023 workshop is the fourth edition of the successful Integrity Workshop series [4–6], always held in conjunction with the ACM Conference on Web Search and Data Mining (WSDM). The workshop comprises of a combination of invited talks by reputed members of the Integrity community from both academia and industry and peer-reviewed contributed talks and posters solicited via an open call-for-papers.

2 ORGANIZERS

Lluis Garcia-Pueyo, Meta, is an Engineering Manager at Meta, where he leads the FB App Integrity Distribution pillar focusing on personalization, discovery and reduction of negative experiences in the Facebook App (Feed, Reels, Watch, Search, Stories, Comments, etc).

Panayiotis Tsaparas is Associate Professor at the Department of Computer Science and Engineering at University of Ioannina. His research interests include Data Mining & Machine Learning, Social Network Analysis, and Algorithmic Fairness.

Prathyusha Senthil Kumar, Meta, is an ML Engineering Manager at Meta, where she leads the Facebook Integrity Ranking & Controls ML team, responsible for making Facebook App a safe, trustworthy & uplifting place for users by using cutting edge machine learning techniques. Her team builds scalable ML models for integrity content understanding, personalized ranking and algorithmic enforcements.

Timos Sellis is the Director of the "Archimedes" Research Unit of the "Athena" Research Center. His research interests include social network data management, streaming data, graph data management, and spatio-temporal database systems.

Paolo Papotti is an Associate Professor at EURECOM, France. His research is focused on data integration and information quality with a focus on computational fact-checking with structured data.

Sibel Adalı is a Professor of Computer Science and Associate Dean of Research and Graduate Studies in the School of Science at Rensselaer Polytechnic Institute. Adalı's work largely focuses on modeling and processing of non-traditional data types in databases, in particular analysis of social network data, computational models of social and information trust and distrust.

Giuseppe Manco is director of research at the institute of High Performance Computing and Networks of the Italian National Research Council (ICAR-CNR) and contract professor at the University of Calabria. His current research interests include machine learning, knowledge discovery and data mining, probabilistic modeling and recommender systems, social network analysis and information diffusion, AI and cybersecurity.

Tudor Trufinescu is a principal engineer working on applied AI for content moderation and platform integrity at Meta. His work also includes the AI Infrastructure for feature engineering, inference and data privacy, and the Facebook search infrastructure for posts, videos, Instagram, marketplace, and places.

Gireeja Ranade is Assistant Teaching Professor in EECS at UC Berkeley. Her research interests include Understanding Misinformation, Information and Control Theory, Protocol Design for wireless control and Safety and Control.

James Verbus is a Senior Staff Machine Learning Engineer on the Anti-Abuse AI Team at LinkedIn. His current focus areas include improving AI developer productivity and the development of advanced machine learning techniques to scalably prevent abuse.

Mehmet Tek is an Engineering Manager at Google, where he leads the Ads Content and Targeting Safety teams. His teams are focused on developing accurate, robust, and fair ML models that understand multi-modal ads content to enforce global Google Ads policies on a wide spectrum of topics, like hate speech, clickbait, gambling, housing, employment, and credit.

Anthony McCosker is Professor of media and communication, Deputy Director of the Swinburne University's Social Innovation Research institute, and Chief Investigator in the ARC Centre of Excellence for Automated Decision Making and society. He researches the impact and uses of social media, data and new communication technologies, with a focus on mental health, digital citizenship, inclusion and literacy.

REFERENCES

[1] Aws Albarghouthi and Samuel Vinitsky. 2019. Fairness-Aware Programming. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Atlanta, GA, USA) *(FAT* '19)*. Association for Computing Machinery, New York, NY, USA, 211–219. https://doi.org/10.1145/3287560.3287588

[2] Guillaume Bouchard, Guido Caldarelli, and Vassilis Plachouras. 2019. ROME 2019: Workshop on Reducing Online Misinformation Exposure. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval* (Paris, France) *(SIGIR'19)*. Association for Computing Machinery, New York, NY, USA, 1426–1428. https://doi.org/10.1145/3331184.3331645

[3] Lluís Garcia-Pueyo, Anand Bhaskar, Prathyusha Senthil Kumar, Panayiotis Tsaparas, Kiran Garimella, Yu Sun, and Francesco Bonchi. 2021. MISINFO 2021: Workshop on Misinformation Integrity in Social Networks. (2021).

[4] Lluís Garcia-Pueyo, Anand Bhaskar, Panayiotis Tsaparas, Aristides Gionis, Tina Eliassi-Rad, Maria Daltayanni, Yu Sun, and Panagiotis Papadimitriou. 2020. Integrity 2020: Integrity in Social Networks and Media. In *Proceedings of the 13th International Conference on Web Search and Data Mining* (Houston, TX, USA) *(WSDM '20)*. Association for Computing Machinery, New York, NY, USA, 905–906. https://doi.org/10.1145/3336191.3371880

[5] Lluís Garcia-Pueyo, Anand Bhaskar, Roelof van Zwol, Timos Sellis, Gireeja Ranade, Prathyusha Senthil Kumar, Yu Sun, and Joy Zhang. 2021. Integrity 2021: Integrity in Social Networks and Media. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining* (Virtual Event, Israel) *(WSDM '21)*. Association for Computing Machinery, New York, NY, USA, 1159–1160. https://doi.org/10.1145/3437963.3441841

[6] Lluís Garcia-Pueyo, Panayiotis Tsaparas, Anand Bhaskar, Prathyusha Senthil Kumar, Roelof van Zwol, Timos Sellis, Anthony McCosker, and Paolo Papotti. 2022. Integrity 2022: Integrity in Social Networks and Media. https://doi.org/10.48550/ARXIV.2209.11867

[7] Richard Han and Neil Shah. 2019. Cybersafety 2019: The 4th Workshop on Computational Methods in Online Misbehavior. In *Companion Proceedings of The 2019 World Wide Web Conference* (San Francisco, USA) *(WWW '19)*. Association for Computing Machinery, New York, NY, USA, 146–147. https://doi.org/10.1145/3308560.3316493