

# Reports of the Workshops Held at the 2020 International Association for the Advancement of Artificial Intelligence Conference on Web and Social Media

*Tarek Abdelzaher, Jisun An,  
Marya Bazzi, Jeremy Blackburn, Ugur Kursuncu, Yelena Mejova,  
J r mie Rappaz, Horacio Saggion, Panayiotis Smeros, Sanjaya Wijeratne, Ning Yu*

■ *The workshop program of the Association for the Advancement of Artificial Intelligence's Fourteenth International Conference on Web and Social Media was held June 8 to 20, 2020. The conference venue, which had originally been Atlanta, Georgia, USA, had to be revamped into a virtual edition, due to the 2019 Corona Virus Disease pandemic. There were five full-day workshops in the program. They included *Emoji Understanding and Applications in Social Media (W1)*, *Social Sensing: Narrative Analysis on Social Media (W2)*, *News and Public Opinion (W3)*, *Cyber Social Threats (W4)*, and *Mediate: Social and News Media Misinformation (W5)*. This report contains summaries of all five workshops.*

## Emoji Understanding and Applications in Social Media

The third International Workshop on Emoji Understanding and Applications in Social Media (ICWSM)<sup>1</sup> was held on the eighth of June 2020 as a half-day online meeting. This workshop brought together computer and social science researchers as well as leading industry practitioners to discuss and exchange ideas on understanding social, cultural, communicative, and linguistic roles of emoji while leading the discussions on building novel computational methods to understand them.

With the rise of social media, emoji have become an extremely popular way to enhance electronic communication. Social media data has been used to study how emoji are used across different languages, cultures, and user communities, and as features to learn machine learning models to solve problems spanning many applications. The ability to automatically process, derive meaning, and interpret text fused with emoji is essential as society embraces emoji as a standard form of online communication. However, the pictorial nature of emoji — that (the same) emoji may be used in different contexts to express different meanings, that the same emoji may be rendered differently across different devices and platforms, and that emoji are used in different cultures and communities over the world who interpret emoji differently — make it difficult to apply traditional natural language processing techniques to analyze and understand them. The goal of this workshop was to stimulate research and discussion on developing novel approaches to address these challenges.

The workshop program consisted of a keynote speech by Francesco Barbieri (Research Scientist, Snap Inc.), who presented his talk, *Why do Machines Struggle to Understand Emoji?* where he discussed the challenges in developing computational models to represent the meaning of emoji. He emphasized how emoji interpretation could be influenced by different cultures, user groups, geographies, and seasonality, thus, making it extremely difficult to develop generalizable emoji interpretation models. He also discussed biases in emoji use and the challenges in building personalized emoji models.

Jennifer Lee (co-founder, Emojination and Vice Chair, Unicode Emoji Subcommittee) gave an invited talk on The Unicode Consortium's<sup>2</sup> emoji approval process, highlighting the gaps in current emoji categories.

The five research papers presented at the workshop covered a wide variety of topics including emoji-embedding models, memetic sticker adoption, emoji-word-order relations, and emoji sentiment analysis. Several papers discussed emoji research problems that are of interest to the computational social science researchers and linguists such as analyzing Animoji adoption across genders, memetic sticker use in Chinese social media platforms, and examining the word-order relations in emoji sequences. Other papers were focused on computer science research problems, such as examining similarities between emoji and word-embedding models and how emoji interpretation and sentiment varies across different emoji vendor platforms. Our interdisciplinary workshop program sparked highly engaging and thought-provoking discussions among nearly fifty researchers who attended the workshop.

To minimize the effect of not having an opportunity to network in-person due to the online nature of this year's meeting, an hour-long virtual networking event — which attracted lively participation — was conducted at the end of the workshop. During

the networking event, attendees had two minutes each to share their research interests with the others followed by a question-and-answer session where research ideas were discussed in detail. A document containing attendees' contact details along with their research interests was disseminated at the end of the workshop to encourage future collaborations among attendees.

Sanjaya Wijeratne, Horacio Saggion, and Amit Sheth served as co-chairs of the workshop. The accepted papers were posted at the ICWSM workshop proceedings site.<sup>3</sup>

The co-chairs thank Emojination for partnering with Adobe Inc. to sponsor twenty students with their workshop registration fees.

## Social Sensing

The Fifth International Workshop on Social Sensing, SocialSens 2020,<sup>4</sup> was held virtually on June 8, 2020, with a focus on human narrative. The goal of this workshop was to generate insights on online narrative analysis that draw on the best of multiple worlds, from physical signal processing to social science.

The Social Sensing workshop (started in 2015) is a multidisciplinary meeting place that brings together social scientists and computer scientists, interested in social media analysis, around research that interprets social media as measurement instruments. Social media democratized information production offering an unprecedented view into human habits, customs, culture, stances, and indeed descriptions of physical events that transpire in the world. They also give unprecedented opportunities to spread misinformation, influence opinion, distract from truth, or advance specific agendas, hidden or overt. Traditional sensors in physical domains respond to signals with well-understood behaviors and propagation models. What can one learn from physical signal processing literature to enable novel social media analysis methods? This scope brings about new interdisciplinary research challenges and opportunities at the intersection of communication and sensing, social network analysis, information theory, data mining, natural language processing, artificial intelligence, and social science. With a focus on narrative analysis, the Social Sensing workshop brought together over 100 researchers from academia, industry, and government to share recent advances in both theoretical and experimental research to answer questions such as how can a narrative be described and captured with computational techniques? How can one separate different narratives (or spins) on the same topic? How can one infer social attitudes from narrative attributes? And ultimately, can one predict the evolution of narratives?

Defining online narrative in a way that is relevant to researchers from different backgrounds is challenging. Nonetheless, this workshop distilled important elements from online narratives: influence intent, the moral value of the target population, impact on

behavior, and so on. Acknowledging that narrative is always more than the sum of the parts and the deeper context is important, the workshop participants did not come up with a universal definition and agreed that it is likely to be application-dependent. For example, in the context of simulation of online narrative spread related to the Venezuelan presidential crisis, Jeremy Blackburn (Leidos) defined narrative as “a collection of statements expressing a point of view on some topic. These statements may explain events, interpret the motives of actors, or emphasize specific perspectives.”

Keynote given by Professor Heng Ji from University of Illinois at Urbana-Champaign presented her new research direction on event-centric knowledge base construction from multimedia multilingual sources. Specifically, several novel methods such as inducing hierarchical narrative graph schemas and applying them to enhance various natural language processing tasks to natural language generation were especially well received. A few presenters also talked about their technical exploration of narrative analysis. Allison Koenecke and Jordi Feliu-Faba from Stanford University modeled polarized views on climate change in response to natural disaster events, which is a nice attempt to bring real-world context into online narrative analysis. George Mohler from Indiana University and Purdue University discussed work applying a network Hawkes binomial topic model to track evolving subtopics around the 2019 Corona Virus Disease (COVID-19) subtopics. Leidos’ work on developing narrative annotation guidelines and a supervised narrative labeler led to discussions on how to interpret and improve low interannotator agreement for such a highly context-dependent annotation task. This suggests an important subarea for further research.

Panelists in the Operational Value of Narrative Research panel felt that the cognitive and behavioral impact of narratives is what makes it important. Panelists also agreed on the importance of framing and building cognitive profiles for the target population. Brian Kettler from the Defense Advanced Research Projects Agency pointed out that studying online narratives requires interdisciplinary collaboration from the beginning: social, computing, engineering, marketing, and so on. Edward Palazzolo from the Army Research Office emphasized the importance of narrative context, raising concerns about temporal validity given that populations often respond differently when an event is happening now versus six months later. Jonathan Pfautz from Leidos suggested focusing on observable and small narrative components first, as humans are complex. Bill Casebeer, former Defense Advanced Research Projects Agency Program Manager for the Narrative Networks program discussed detecting and connecting semantic and structural primitives to psychologic mechanisms. He also suggested researchers to build narrative research on top of existing theories and infrastructures, such as behavioral economics and information foraging.

Online narrative analysis is an evolving area of research. A few pressing challenges were identified at the end of the workshop: how to decide the best level of narrative representation and the best practices for developing ground truth; how to simplify user behaviors; and how to effectively respond to or counter a malicious narrative.

Jiawei Han and Ning Yu served as the general chairs and Emilio Ferrara and Tarek Abdelzaher served as the program chairs for the workshop. This report was written by Ning Yu and Tarek Abdelzaher.

## News and Public Opinion

The fourth international workshop on News and Public Opinion, NECO 2020,<sup>5</sup> was held virtually during the morning of June eighth, in conjunction with the 2020 International Association for the Advancement of Artificial Intelligence Conference on Web and Social Media. Following the last series of successful workshops, the workshop’s goal was to gather scholars from the fields of social science and computer science together and explore different approaches to tackling various issues in media studies using large-scale data. Three invited talks and three contributed talks covering multiple disciplines were presented, and the audience actively participated in the discussions through the Zoom platform.

The role of news media becomes more and more critical, based on its vast reach through social media platforms. It is well known that what is reported by the news media (and how it reports the incidents) is crucial to establishing a public understanding of, and attitudes toward, the incidents. In this sense, the potential bias of news coverage and its impact on public opinion has been actively studied in recent years. The fourth international workshop on News and Public Opinion was intended to build better understandings of news media, public opinion, and their interactions under this situation.

The primary theme of the invited and contributed talks presented at the News and Public Opinion workshop is aligned with these lines of research trends. First, regarding the diversity and fairness of news coverage and the audience, Giovanni Luca Ciampaglia (University of South Florida) gave an invited talk about a data-driven analysis of the audience diversity and algorithmic bias in news recommendation. Eni Mustafaraj (Wellesley College) presented a disproportionate amplification of far-right news sources for election news coverage in Google’s top stories.

Second, regarding the agenda-setting role of news media, Laszlo Horvath (University of Exeter) presented how quantitative analyses can be combined with qualitative methods (for example, ethnography) to bring a more comprehensive understanding of the media’s agenda-setting and public’s news sharing. Masha Krupenkin (Boston College) demonstrated the media’s power as a disseminator of public health messaging and a mediator of presidential agenda-setting

through the analysis of COVID-related news coverage and web-search data.

Third, regarding understanding the public, Indira Sen (Gesellschaft Sozialwissenschaftlicher Infrastruktureinrichtungen eV [German Social Science Infrastructure Services], Leibniz Institute for the Social Sciences) showed how the public opinion could be better measured on social media. Dean Freemon (University of North Carolina at Chapel Hill) presented the analysis of 86 million tweets posted in 2017 to examine how users from across the political spectrum engage differently with news issues and major media outlets on Twitter.

As was done at previous News and Public Opinion workshops, the 2020 workshop offered an excellent venue for participants to share their expertise and knowledge on news and public opinion. The participants enjoyed the virtual workshop and expressed vital interests in future workshops.

Jisun An (a scientist at Qatar Computing Research Institute, Hamad Bin Khalifa University), Haewoon Kwak (a senior scientist at Qatar Computing Research Institute, Hamad Bin Khalifa University), and Fabrício Benevenuto (an associate professor at Universidade Federal de Minas Gerais) served as co-chairs of this workshop. The workshop was part of the 2020 International Association for the Advancement of Artificial Intelligence Conference on Web and Social Media.<sup>6</sup> This report was written by Jisun An.

## Cyber Social Threats: Challenges, with Methodological and Ethical Considerations

The Cyber Social Threats Workshop<sup>7</sup> was held on the morning of June 8, 2020. The goal of this workshop was to investigate the challenges and novel approaches in the analysis of online harmful communications, concerning social, cultural, emotional, communicative, and linguistic aspects.

In recent years, online platforms have been misused for promoting harmful content and behavior such as extremism, harassment, fake news, misinformation, human trafficking, and gender-based violence, among others affecting our society. Such content and behaviors are inherently complex, making the recognition of their narratives challenging for researchers as well as social media companies. In the pursuit of stimulating novel research directions, exchanging ideas and experiences and identifying new opportunities for collaboration, the Cyber Social Threats workshop provided a forum to bring together researchers and practitioners from both academia and industry in the areas of computational social sciences, social network analysis and mining, natural language processing, computational linguistics, human-computer interaction, and cognitive scientists to present their related, fundamental research and emerging applications.

The two invited keynotes provided diverse insight into both academic and industrial perspectives on dealing with potentially harmful behaviors. First, Alexandra Olteanu, Microsoft Research, presented Challenges to Measuring Objectionable Behavior Online by Humans and Machines, focusing on the lack of understanding in the implications of computational models on online harmful behaviors, such as inadvertently reinforcing and amplifying them. The second speaker, Mikey Cohen from the Network Contagion Research Institute, shared his motivation behind an unexpected career change from being a leader at Netflix to fighting hate online. He described the challenges to deciphering and resolving the seemingly insurmountable and ever-changing problems online and how a collective effort between academia and the industry in Silicon Valley would address them.

The workshop had eight accepted papers, four of which were related to the pressing issue of the COVID-19 pandemic and related misinformation efforts. The COVID-19 paper session included presentations on the prevalence of low-credibility information, the impact of misinformation on conversations in Telegram chat groups, the role of Twitter bots in the spread of misinformation in the Middle East, and a multilingual news dataset. Four other papers were presented: methodological and ethical considerations in the study of toxic behaviors; a social media analysis tool; fake news detection; and a study on reactions to a tragedy on social media. This was a surprisingly diverse, yet interestingly well-connected set of topics. In particular, starting this session with the paper discussing ethical considerations led to some lively discussion as the authors of quantitative and systems papers in the session had experienced many of the concerns and decisions that were well articulated by the more theory-oriented work. Finding this type of interdisciplinary common ground was a core goal of the workshop.

All papers were reviewed by at least three multidisciplinary program committee members (twenty-one in total) in the fields of computer science, information science, political science, sociology, psychology, and communications. The best paper was selected by a committee comprised of four program committee members, and the authors were invited to submit an extended version of their paper to a special issue on Cyber-Social Health: Promoting Good and Countering Harm on Social Media, for the Institute of Electrical and Electronics Engineers' International Conference on Collaboration and Internet Computing.<sup>8</sup>

Finally, a synthesis exercise session took place where the participants brainstormed ideas that they found most important, urgent, and high-impact for potential future work and collaborations. The ideas included efforts to mitigate doxing (publishing private information about individuals online) and mental health problems induced by COVID-19, as well as moves toward greater transparency (such as methods

for bot detection), evaluation of therapies for those affected by hate speech, and discussion of the actionable application of research to policy changes.

At the end of the workshop, the participants expressed their interest in collaborating on the identified problems and areas, as well as participating in future workshops to be organized in the coming years. Ugur Kursuncu, Yelena Mejova, Jeremy Blackburn, and Amit Sheth served as co-chairs of this workshop. This report was written by Ugur Kursuncu, Yelena Mejova, and Jeremy Blackburn. The papers were posted at the ICWSM workshop proceedings site.<sup>9</sup>

## Mediate: Social and News Media Misinformation

The first Mediate workshop (MEDIATE Workshop 2020)<sup>10</sup> was held virtually on the afternoon of June eighth, 2020, as part of the ICWSM. The main goal of the workshop was to bring together media practitioners and technologists to discuss new opportunities and obstacles that arise in the modern era of information diffusion. This year's theme was social and news media misinformation, the various forms it can take, and the different ways one can approach it.

Social and news media misinformation is a pervasive problem of our time, which heavily permeates our day-to-day news consumption. The Mediate workshop brought together academics and practitioners, spanning many different backgrounds, from computer science, natural language processing, and network science to journalism and education. The goal of the workshop was to consider the different streams of research and work on misinformation, and to highlight some important open questions.

The first topic of our workshop was social initiatives against misinformation and consisted of two invited talks. It opened with Amy X. Zhang (University of Washington) discussing how people with varying expertise judge news credibility and reliability. In their work, they collaborated with Credibility Coalition,<sup>11</sup> a research community that fosters a collaborative approach to understanding the veracity, quality, and credibility of online information. The key take-away of the talk was that people's ratings have higher correlation to journalism experts compared with science experts as well as that raters' gender and political leaning impact their ratings. Juliane von Reppert-Bismarck, founder and Chief Executive Officer of Lie Detectors,<sup>12</sup> then spoke about enabling schoolchildren to identify misinformation online. Some take-aways were that young people are now more acutely aware of the danger of misinformation because of COVID-19; that technology currently plays a limited role in increasing their media literacy; and that trust is an important issue for them.

The second topic of the workshop was news analytics and consisted of three contributed talks. Antonia

Saravanou (National and Kapodistrian University of Athens) presented a learning to rank approach for detecting notable news stories. Elena Kochkina (University of Warwick and The Alan Turing Institute) discussed the relationship between the tasks of rumor, stance, and veracity classification in social media conversations. Finally, Luca Luceri (University of Applied Sciences and Arts of Southern Switzerland) presented a study on social media manipulation campaigns, where they showed that bots strategically mimic the human temporal activity, and balance their interaction among the human population and bots.

The third topic was political and scientific misinformation and consisted of three invited talks. It started with Sibel Adali (Rensselaer Polytechnic Institute) who described how social trust is not a single concept but incorporates many different constructs relating to both competence and integrity of individuals. These constructs may be evaluated differently in a cognitive sense, and models that incorporate this distinction are able to provide a nuanced solution to important problems. Filippo Menczer (Indiana University) presented four reasons why social networks make us vulnerable to misinformation. First, echo chambers make users exhibiting specific characteristics (for example, political views) easy to target. Second, viral dynamics could lead to high content exposure. Third, engagement bias; that is, users are likely to share content that received significant engagement. Fourth, social bots are commonly used in manipulation campaigns. Ivan Oransky (Medscape and New York University), co-founder of Retraction Watch,<sup>13</sup> discussed misinformation in scientific publications that manifests as retractions. The number of retractions is significant, and the rate of retraction seems to show an increasing trend over several years. The causes for retractions can vary widely (for example, error; fake peer review; plagiarism; not reproducible).

Talks in our workshop focused on detecting misinformation in its many forms, on understanding its different facets and what enables it to prosper, and on ways to mitigate it (for example, through media literacy). Two high-level themes, trust and intention, have emerged as open questions. How can one quantify or instill trust in media? How can one causally link, in an automated way, the presence of misinformation with human intention or action? Multi-disciplinary workshops on misinformation, or more generally on digital media, will hopefully continue to bring together the wide range of efforts in this space and highlight the challenges and opportunities ahead.

Panayiotis Smeros, Jérémie Rappaz, and Marya Bazzi served as co-chairs of this workshop. The contributed papers of the workshop were posted at the ICWSM workshop proceedings site,<sup>14</sup> and all talks are publicly available online.<sup>15</sup>

This report was written by Jérémie Rappaz, Panayiotis Smeros, and Marya Bazzi.



## ICWSM-21

The Fifteenth International AAAI Conference on Web and Social Media will be held at the Georgia Tech Hotel and Conference Center in Atlanta, Georgia USA from June 7–10, 2021.

[www.icwsm.org](http://www.icwsm.org)

### Notes

1. [workshop-proceedings.icwsm.org](http://workshop-proceedings.icwsm.org).
2. [home.unicode.org](http://home.unicode.org).
3. [icwsm.org/virtual/2020/workshops\\_5.html](http://icwsm.org/virtual/2020/workshops_5.html).
4. [socialsens.web.illinois.edu/](http://socialsens.web.illinois.edu/).
5. [neco.io/](http://neco.io/).
6. [www.icwsm.org/2020/index.html](http://www.icwsm.org/2020/index.html).
7. [cysoc.aiisc.ai/](http://cysoc.aiisc.ai/).
8. [www.x-mol.com/paper/1241038306465484800](http://www.x-mol.com/paper/1241038306465484800).
9. [workshop-proceedings.icwsm.org](http://workshop-proceedings.icwsm.org).
10. [mediateworkshop.github.io/](http://mediateworkshop.github.io/).
11. [credibilitycoalition.org/](http://credibilitycoalition.org/).
12. [lie-detectors.org](http://lie-detectors.org).
13. [retractionwatch.com](http://retractionwatch.com).
14. [workshop-proceedings.icwsm.org](http://workshop-proceedings.icwsm.org).
15. [icwsm.org/virtual/2020/workshops\\_5.html](http://icwsm.org/virtual/2020/workshops_5.html).

**Tarek Abdelzاهر** is a professor and Willett Faculty Scholar in the Department of Computer Science at the University of Illinois at Urbana-Champaign, Urbana-Champaign, Illinois.

**Jisun An** is a scientist at the Qatar Computing Research Institute, Hamad Bin Khalifa University, Ar-Rayyan, Qatar.

**Marya Bazzi** is a Turing Research Fellow at the Alan Turing Institute, London, United Kingdom.

**Jeremy Blackburn** is an assistant professor in the Department of Computer Science at Binghamton University, Binghamton, New York.

**Ugur Kursuncu** is a post-doctoral fellow at the Artificial Intelligence Institute, University of South Carolina, Columbia, South Carolina.

**Yelena Mejova** is a research leader at the Institute for Scientific Interchange Foundation, Turin, Italy.

**J r mie Rappaz** is a senior PhD student at  cole Polytechnique F d rale de Lausanne, Lausanne, Switzerland.

**Horacio Saggion** is a professor and the head of the Large-Scale Text Understanding Systems Lab at the Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain.

**Panayiotis Smeros** is a senior PhD student at  cole Polytechnique F d rale de Lausanne, Lausanne, Switzerland.

**Sanjaya Wijeratne** is a research scientist at Holler Technologies, Inc., San Mateo, California.

**Ning Yu** is a senior research scientist of the Artificial Intelligence/Machine Learning Accelerator at Leidos, Reston, Virginia.