

# Reports of the Workshops Held at the 2018 International AAAI Conference on Web and Social Media

*Jisun An, Rumi Chunara, David J. Crandall, Darian Frajberg, Megan French,  
Bernard J. Jansen, Juhi Kulshrestha, Yelena Mejova, Daniel M. Romero,  
Joni Salminen, Amit Sharma, Amit Sheth, Chenhao Tan,  
Samuel Hardman Taylor, Sanjaya Wijeratne*

■ *The workshop program of the Association for the Advancement of Artificial Intelligence's 12th International Conference on Web and Social Media (AAAI-18) was held at Stanford University, Stanford, California, June 25, 2018. There were 14 workshops in the program. This report contains summaries of those workshops.*

The workshop program of the Association for the Advancement of Artificial Intelligence's 12th International Conference on Web and Social Media (AAAI-18) was held at Stanford University, in Stanford, California, on June 25, 2018. There were 14 workshops in the program: Algorithmic Personalization and News: Risks and Opportunities; Beyond Online Data: Tackling Challenging Social Science Questions; Bridging the Gaps: Social Media, Use and Well-Being; Chatbot; Data-Driven Personas and Human-Driven Analytics: Automating Customer Insights in the Era of Social Media; Designed Data for Bridging the Lab and the Field: Tools, Methods, and Challenges in Social Media Experiments; Emoji Understanding and Applications in Social Media; Event Analytics Using Social Media Data; Exploring Ethical Trade-Offs in Social Media Research; Making Sense of Online Data for Population Research; News and Public Opinion; Social Media and Health: A Focus on Methods for Linking Online and Offline Data; Social Web for Environmental and Ecological Monitoring and The ICWSM Science Slam.

The workshops were held on the first day of the conference. Workshop participants met and discussed issues with a selected focus — providing an informal setting for active exchange among researchers, developers, and users on topics of current interest. Organizers from nine of the workshops submitted reports, which are reproduced in this report. Brief summaries of the other five workshops have been reproduced from their website descriptions.

## Algorithmic Personalization and News: Risks and Opportunities

The Algorithmic Personalization and News workshop aimed at broadening the debate on algorithms and news to discuss the risks and opportunities of personalization for news consumption and to deliberate suitable methods for auditing algorithmic personalization and its impact on users' news consumption behaviors.

Algorithmic systems have become ubiquitous in our modern lives, and they exert a significant influence on many aspects of our daily routines. This influence includes shaping the news and information that we are exposed to in online news environments through information retrieval algorithms like recommendation systems (for example, Google News or Facebook News Feed), as well as search systems on the web or social media platforms. These algorithms alter the news that users consume. They can be deployed directly on news websites in the form of recommendations, or they can impact news consumption via algorithmic social media feeds or search engines by deciding what news users see online.

The workshop was intended to provide a platform to discuss the risks and opportunities of algorithmic personalization systems in the online news environments, and also to encourage the design of auditing mechanisms for such news personalization algorithms. Our goal was to foster dialogue between the different communities interested in the topic, and we achieved this aim by bringing together participants from academe in the fields of computer science, media studies, journalism, and communication sciences with industry practitioners from mainstream media, social media organizations, and startups.

Christo Wilson (Northeastern University) kicked off the workshop program with the keynote talk "Google Search and Politics," in which he discussed his work on auditing different algorithmic components of Google search in the context of political information seeking. The workshop also featured a moderated panel focused on the two central themes of risks and opportunities of algorithmic personalization for news, with members from academe as well as industry, comprising of Jisun An (scientist at Qatar Computing Research Institute, HBKU), Rich

Jaroslovsky (vice president for content and chief journalist, SmartNews Inc.), and Titus Plattner (John S. Knight Journalism Fellow at Stanford University).

The workshop also featured short talks on the accepted abstracts. The first theme of work presented concentrated on the risks of news personalization and covered a wide variety of topics, including a case study of Google News personalization, the privacy implications of personalization persuasion architectures, and the study and mitigation of user input bias in personalization platforms. The other prominent theme focused on the opportunities that personalization platforms afford for news production, dissemination, and consumption. Participants presented techniques for increasing user engagement for diverse news formats ranging from videos of UK parliamentary debates to industry experiences by using personalized news recommendations and summaries for textual articles and radio snippets. Also presented were factors for article-level personalization of the content for different types of users.

With a varied audience representing both industry and academe, from across different fields of study, the workshop turned out to be the perfect venue for discussing the algorithmic personalization of news. The consensus among the participants was that while it is important to audit and study the impact that personalization frameworks have on different aspects of news production, dissemination, and consumption, it is also equally vital to make optimal use of these technologies to improve the news generation and dispersion processes.

Aniko Hannak (Center for Network Science, Central European University), Nicholas Diakopoulos (School of Communication, Northwestern University), Juhi Kulshrestha (CSS Department, GESIS - Leibniz Institute for the Social Sciences), and Cornelius Puschmann (Hans Bredow Institute for Media Research, University of Hamburg) served as the cochairs for the workshop. This report was written by Juhi Kulshrestha.

## Beyond Online Data: Tackling Important Social Science Challenges

The Beyond Online Data workshop brought together social scientists and computer scientists to think about new ways to utilize data for addressing challenging social science questions. The goal of the workshop was to promote question-driven research and advance work in the nascent field of computational social science.

The abundance of online social data has provided a new lens for asking social science questions on issues such as interpersonal relationships, information diffusion, persuasion, and language use. However, most social processes thrive at the intersections of online and offline worlds. It thus becomes necessary to connect online studies to the offline world, espe-



cially for tackling broad social science questions such as information access, education, healthcare, migration, discrimination, and poverty. In this workshop, we discussed novel ways of utilizing and combining offline and online data, and creating new data through observation and experiments. Instead of starting with available datasets, the focus was on starting with a substantive question and then leveraging diverse potential datasets, including both online data and offline data.

The workshop brought together researchers from different fields of computer science and social sciences, including economics, sociology, and communications. There were five invited talks and four contributed talks. The talks represented a diverse and challenging set of real-world problems solved using various sources of online and offline data, including fitness and exercise, science, social media, news consumption, demographics, healthcare, advertisement, and organizational data. The following themes emerged: (1) the use of machine learning combined with rich datasets to gain insights for decision-making in healthcare, (2) the use of causal inference tech-

niques and natural experiments to establish causal relationships in cases where prediction is not sufficient to gain a full understanding of the problem, and (3) the use of large-scale online and offline activity data to promote well-being and improve lives.

The application of computational techniques in societally critical domains also raises ethical concerns, many of which were brought out in a panel discussion with the invited speakers. The panel also discussed issues around data availability, particularly in cases where proprietary data can only be accessed by few researchers. The panel identified the strengths and potential shortcomings of solutions such as the Social Science One initiative by Facebook and academics from multiple universities, which enables partnerships between academic researchers and industry to share data that can benefit society in an open, fair, and responsible way.

The workshop was organized by Michael Macy, Daniel M. Romero, Amit Sharma, and Chenhao Tan.

This report was written by Daniel Romero, Amit Sharma, and Chenhao Tan.

## Bridging the Gaps: Social Media Use and Well-Being

The goal of the Bridging the Gaps workshop was to identify difficult and pressing issues facing academic and industry scholars studying the effects of social media use on well-being. The topic of social media and well-being is of growing interest to academic researchers, industry professionals, and platform designers. Despite broad-reaching interests in the topic, scholars still have a limited understanding of why social media influences well-being. The overarching goal of this workshop was to map the open questions of social media and well-being, identify the processes underlying the relationship between social media and well-being, and develop potential design solutions.

This workshop brought together a diverse group of researchers and professionals from a variety of backgrounds to discuss the emerging trends of social media and well-being, including communication, information science, clinical and social psychology, human development, and computer science. The themes of the workshop emerged from small and large group discussion where participants brainstormed solutions for issues across a variety of domains, including theory, methods, findings, and dissemination. The second section of the workshop involved a design sprint to prototype design solutions to better support well-being through social media.

One theme that emerged across groups was the need for more theoretical grounding of research on this topic. Although studies often acknowledge theories of well-being, the contribution to understanding of well-being is often underdeveloped. There is a clear need to consider the underpinnings of social media technologies, such as affordances, to help research transcend the fast-paced changes of social media platforms. One solution offered was for researchers to be intentional in the types of well-being they investigate and to test related theoretical processes. In addition, reviewers can help by rewarding work on the mechanisms underlying social media effects.

A second research theme was the methodological limitations of the current literature. Several groups noted that correlation is king and that there is little advancement into longitudinal and experimental studies. Furthermore, most research tends to focus on one specific platform, without considering the larger media context, including offline interactions. Other groups pointed out that there is unequal access to data on social media behaviors and that this disjunction prevents some researchers from being able to engage with critical questions. A solution to these methodological challenges is to create more channels and incentives for collaborations between industry and academic researchers.

Building from the themes established at the beginning of the workshop, teams worked to design social media that supports well-being. The design themes included blending offline and online, mental health, intentionality of use, community and belonging, and relational ties. Several of the design solutions focused on designing trusted connections who might provide assistance with a stressful event or life transition or building community through charitable giving. Other designs playfully thought through how social media companies could design to support casual conversations between friends offline using machine-learning. One notable outcome of this design sprint was translating these large-scale research problems that had been identified into actionable changes to social media platforms. The design activity highlighted that while many research issues remain, there are opportunities to put our cumulative knowledge on well-being and social media into practice.

Another critical issue raised throughout the workshop is the difficulties researchers face in communicating both the positive and negative effects of social media on well-being. Although there is a wealth of knowledge about social media and well-being, public discussion of the topic often emphasizes negative effects. The group of researchers in the room agreed that more training on how to translate research to the general public is needed to make research on this topic more impactful. However, many noted that thoughtful science communication is time-consuming and rarely incentivized.

Overall, the workshop discussion illustrated the need to add more nuance to theoretical and media discussions of the topic, as well as the need to create avenues for more collaborations with a wider range of methodologies. Participants noted that workshops that bring together diverse perspectives are important for addressing the gaps in the literature and pushing the field forward. Many expressed interest in continuing conversations that arose during the day.

Megan French, Sunny Liu, Eden Litt, and Samuel Hardman Taylor co-organized this workshop on social media use and well-being. This report was written by Samuel Hardman Taylor and Megan French.

## Chatbot

With AI technologies invading every part of our daily life, processing and delivering information has been fundamentally changed, bringing it ever closer to human-style communication. In the AI era, chatbots will become the new interface between humans and information, humans and services, and humans and society. Dialoging is one of the most common and intuitive ways for humans to communicate. Previously, the web interface of searching and browsing provided the main vehicle of exchange between humans and online information, but this unidirectional method of inquiry will be soon replaced by

chatbots. We envision that the chatbot will become the front-end interface of almost everything digital.

Developing better chatbot technology requires interdisciplinary attention. The backend technologies of chatbots rely upon domain expertise in information/cognitive summary, sentiment analysis, topic modeling, natural language processing, text mining, and machine learning. The front-end technologies are built from a deep understanding of user behavior and engagement, cognitive thinking, UI design, and multimodal interaction design. In addition, designing chatbots entails the complicated social interaction of individuals and groups, which includes peer communication, conformity, social and society pressures, trends, personality, social bonding, and more.

Ying Ding, Bing Liu, James Shanahan, and Jie Tang organized the workshop. No report was submitted to *AI Magazine*.

## Data-Driven Personas and Human-Driven Analytics: Automating Customer Insights in the Era of Social Media

The Data-Driven Personas and Human-Driven Analytics workshop dealt with the use of online analytics data for automating customer insights, specifically via the use of personas. Automation can bring many benefits, including increased speed of producing reports, removing judgment errors, and reducing cost of customer analysis. However, efforts toward automation also involve many challenges, such as data retrieval, presentation, and validation of the usefulness of such systems. We illustrated these problems by presenting a methodology called *automatic persona generation* that summarizes real data on online behaviors and demographics into easily interpretable personas. This system and its key challenges were introduced to the participants, after which we divided into teams, so that each team picked a subproblem to solve.

There were four teams working on the problems. Overall, the participants came up with a rich set of ideas. One team was interested in the usefulness of personas. They came up with an idea of an ad creation experiment to test the effect of personas on real and tangible work output. The transparency of persona profiles was also noted as an issue: from prior research, we know that not including explanations of how information is produced can hinder trust in users and negatively affect their experience of systems that rely on algorithmic filtering. The team concluded that determining how to provide explanations in the context of automatic analytics requires empirical user studies.

Another team worked on the consistency problem of automatic persona profiles, meaning that the

selected information in the persona profile is not always internally coherent (for example, quotes are from a male while the persona is female). Their approach involved classifying comments by age, gender, and location and then matching the appropriate ones to demographic attributes to ensure persona consistency.

Another team was interested in the applicability of social network analysis, that is, graph features, to personas. They postulated that in the absence of direct information, we could learn about the preferences and traits of the persona by analyzing the connections, such as followers, of individuals we have information about.

One team was interested in image generation, or transformation, using generative adversarial networks. The theoretical promise of deep learning could be to transform a set of seed images to closely matching facial pictures with varying age, gender, and even ethnicity. This solution is targeted toward the challenge of finding and downloading thousands of individual images to portray automatically generated personas.

In conclusion, even though the automation of customer analytics has many benefits, there are considerable technical and social challenges in the process. Technical challenges relate to the availability of data and especially the accuracy of inferring various user attributes from user-generated social media content. Moreover, automation has been questioned, as algorithms may pose bias and result in stereotypical or nonmeaningful data portrayals. In our workshop, we explored this duality of data- and human-driven analytics. While data-driven analytics is focusing on the accuracy and availability of data to support decision-making, human-driven analytics is defined as the presentation and analysis of insights about users or customers that highlights qualitative attributes over numbers. Human-driven analytics is based on users' information needs by automatically adapting or giving users interactive choices for information presentation.

The workshop chairs were Jim Jansen, Joni Salminen, Lene Nielsen, and Matti Mäntymäki. This report was written by Joni Salminen and Bernard J. Jansen.

## Designed Data for Bridging the Lab and the Field: Tools, Methods, and Challenges in Social Media Experiments

The Designed Data for Bridging the Lab and the Field workshop explored the methodological middle ground between the two environments of lab and field. Research using repurposed observational data from online platforms has transformed the study of online behavior, and while such data has high external validity, it also presents challenges for establish-

ing causality and replication. By contrast, experimental studies offer high internal validity, discovery of causal relationships, and ease of replication, but the rigid control over the settings and interactions of an experiment can limit generalizability. New methods work toward bridging these two research contexts, either by bringing the lab to the field to recruit more diverse participants in a natural setting or by simulating natural settings and interactions of the field in the lab. This workshop showcased more naturalistic experimental paradigms, innovative tools and methods, and challenges in conducting research to optimize both internal and ecological validity.

The workshop was organized by Dominic DiFranzo, Natalie Bazarova, Shyam Sundar, and Jeff Hancock. No report was submitted to *AI Magazine*.

## Emoji Understanding and Applications in Social Media

The Emoji Understanding and Applications in Social Media workshop brought computer and social science researchers and industry practitioners together to discuss and exchange ideas on understanding the social, cultural, communicative, and linguistic roles of emoji, while leading discussions on building novel computational methods to understand them.

Emoji have become a popular way to enhance electronic communication. They play distinct social and communicative roles compared to other forms of written language. The ability to automatically process, derive meaning, and interpret text infused with emoji is essential as society embraces emoji as a standard form of online communication. Yet a broad variety of reasons make this a challenging task for traditional natural language processing techniques. Key reasons include the pictorial nature of emoji; the fact that (the same) emoji can be used in different contexts to express different meanings; the broad variety of contexts in which a sender may choose to use an emoji (such as the sociocultural aspects and mutual relationships between the sender and the receiver); and the fact that emoji are used and interpreted differently across different cultures and communities around the world. The goal of this workshop was to stimulate research on understanding these challenges in emoji use and on developing novel approaches to analyze, interpret, and understand them.

Tyler Schnoebelen (principal product manager, Integrate.ai) gave the opening keynote speech, “Emoji Are Great and/or They Will Destroy the World.” He discussed the connections of emoji to particular styles and ideologies. Specifically, he compared the way that celebrities, journalists (from right, left, center), and different genders use emoji, while bridging social theories with computational linguistics. Sanjaya Wijeratne, Amit Sheth, and Horacio Saggion presented the tutorial “Improving Emoji Understanding

Tasks Using EmojiNet,” where Wijeratne discussed the way in which knowledge bases of emoji meanings can be used to solve problems of emoji similarity and emoji sense disambiguation.

The eight research papers presented at the workshop covered a wide variety of topics ranging from building and improving emoji lexicons to receiver interpretation of emoji. Several papers discussed emoji research problems that are of interest to the computer science researchers, such as applying temporal variability of emoji usage patterns for improving emoji prediction, learning emoji embedding models via emoji co-occurrence networks, and building emoji lexicons. Other papers were focused on linguistics and social science research problems, such as examining the resemblance of repetitions in emoji use to beat gestures, how emoji are used to express solidarity in social media messages during crisis events, and how receiver interpretation of emoji changes across different genders.

A highly interdisciplinary panel, “The Challenges in Emoji Understanding,” provided an animated and engaging forum to the attendees to discuss the open emoji research problems with leading researchers and practitioners. The panel consisted of Jennifer Lee (vice chair, Unicode Emoji Subcommittee), Keith Winstein (assistant professor, Stanford University), Eric Goldman (professor, Santa Clara University), Rachael Tatman (data scientist, Kaggle), and Francesco Barbieri (postdoctoral researcher, Universitat Pompeu Fabra, Spain). The panel discussion topics included representing emoji using Unicode codepoints, submitting new emoji proposals, challenges in using emoji as a new language, implications of emoji interpretation in the law, and building new computer algorithms to understand emoji.

Sanjaya Wijeratne, Emre Kiciman, Horacio Saggion, and Amit Sheth served as co-chairs of the workshop. The accepted papers are published under CEUR Workshop Proceedings Volume 2130. The workshop program is covered in a *WIRED* article by Arielle Pardes.

The workshop was organized by Sanjaya Wijeratne, Emre Kiciman, Horacio Saggion, and Amit Sheth. This report was written by Sanjaya Wijeratne and Amit Sheth.

## Event Analytics Using Social Media Data

Social media channels enjoy many advantages over traditional media channels, such as ubiquity, mobility, immediacy, and the seamless communication in reporting, covering and sharing real-world events such as the Boston bombings, the NBA finals, and the US presidential elections. Given these advantages, social media posts on Twitter, Facebook, Instagram, WhatsApp and the like can typically reflect events as they happen, in real time. Despite these benefits,

social media channels also tend to be noisy, chaotic, and overwhelming. As a result, the vast amount of noisy social media data poses tremendous challenges for conducting in-depth analysis, which is critical to applications for event playback, journalistic investigation, storytelling, and so on. This workshop was organized to provide a platform to understand the critical challenges in event analytics, to determine the values they bring in various domains, and to enable cross-fertilization of ideas between the different scientific (machine learning, data mining, social computing, natural language processing) and functional (academe and business) groups.

Yuheng Hu and Yu-Ru Lin organized the workshop. No report was submitted to *AI Magazine*.

## Making Sense of Online Data for Population Research

The global spread of social media has generated new opportunities for demographic research, as individuals leave increasing quantities of traces online that can be aggregated and mined for population research. However, whether population research is carried out on traditional data or online data, the key questions remain: What are we aggregating, and how are we aggregating it? To understand population processes requires thoughtful consideration of populations and behaviors as they are manifested in data.

We see growing opportunities for collaboration between computer scientists and social scientists specifically centered on population research methodology. Demography has been a data-driven discipline since its birth, and data collection and the development of formal (that is, mathematical) methods have sustained most of the major advances in our understanding of population processes.

All data come with some type of bias, and an essential responsibility of scientists in any field is to understand the nature of the bias manifested in the data from which they draw inferences and with which they test hypotheses. In this respect, the data generated from human interaction with digital technology is no different from the data traditionally used in social science research, but this data does pose new and interesting problems.

This workshop's major themes were understanding underlying populations and demographic processes, understanding data-generating behaviors, and leveraging scale to ask new questions.

The workshop was organized by Lee Fiorio, Emilio Zagheni, Afra Mashhadi, Bogdan State, and Dennis Feehan. No report was submitted to *AI Magazine*.

## News and Public Opinion

The goal of the third international News and Public Opinion workshop was to gather scholars from the fields of social science and computer science together

to explore different approaches to tackling various issues in media studies using large-scale data. Speakers presented five papers covering multiple disciplines, and the audience actively participated in the discussions.

The media landscape has been changing significantly in the digital era. The internet has weakened the traditional business models, employment structures, coverage practices, and consumption of news. As it becomes easy to take a role as a media outlet, there are hundreds of new online news outlets created and the traditional media outlets compete with them for attention. Along with this rapidly changing media landscape, existing theories need to be revisited and reexamined so that one can have better understanding of the new media environment. Following last year's successful workshop, the goal of this workshop was to seek better approaches to understanding the changes of media landscape and to tracking and tackling the issues in media studies that originated from such changes.

This year, the major theme of the papers presented at the workshop was linguistic characteristics of political discussion online. Nabil Hossain (University of Rochester, New York) proposed a method for detecting creative political slang and showed temporal analysis of the appearance of slang in comments from three different online news sources. Rijul Magu (University of Rochester, New York) built classifiers to predict from the text of the article alone whether uncivil comments would appear for a given news article. The resulting models of the two studies can be widely used for future research on political discussion online. Another major theme, following the theme of last year's workshop, was the news ecosystem. Ben Horne (Rensselaer Polytechnic Institute in Troy, New York) found content republishing patterns among news sources by exploring 92 such sources. He also compared news articles from reliable and unreliable sources and found some universal differences between the two across countries. Lastly, João Vicente Seno Ozawa (University of São Paulo, Brazil) presented a preliminary study on the agenda-setting effects of both mass media and interpersonal communication, conducted by social network analysis.

The keynote address, "News Bots: Agents of Information or Automated Sociopaths?," was given by Nicholas Diakopoulos (School of Communication at Northwestern University). Diakopoulos gave a comprehensive overview of the news bots and their impact on the news ecosystem and social media. The talk prompted an active discussion among researchers from different fields. The group activity that followed was a great place to get to know each other, and each other's research, through one-to-one and group meetings.

The NECO workshop provided an excellent place for participants to share their expertise and knowledge of theories and practice for research on news

and public opinion. The participants enjoyed the workshop and expressed a strong interest in attending a future NECO workshop.

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 Fabrizio Benevenuto (an associate professor at Federal University of Minas Gerais) served as cochairs of this workshop. This report was written by Jisun An.

## Social Media and Health: A Focus on Methods for Linking Online and Offline Data

Focusing on methods for linking online and offline data, the first Social Media and Health workshop aimed at outlining state-of-the-art methods for modeling real-life health data, along with addressing issues of accuracy, representativeness, the ultimate usefulness of social media data, and the privacy and ethical considerations in its use.

Applying web and social media analytics to monitor, study, and perform interventions in the health domain has been an active and ongoing area of research. From Google Flu Trends and its derivatives to Zika virus and vaccination rumor tracking, social media is increasingly used as an alternative or a supplementary signal of health behaviors and beliefs. In this workshop, we addressed the question, How does data from internet sources compare to offline phenomena (collected by traditional means such as medical records or other proxy data sources), at both the individual level and the population level?

The workshop brought together leading practitioners in the fields of health communication and data science, as well as researchers in social media studies, natural language processing, network science, and other related fields. The keynote presentation was delivered by Wen-Ying Sylvia Chou, who is a program director in the Health Communication and Informatics Research Branch (HCIRB) of the Behavioral Research Program (BRP) at the National Cancer Institute at the National Institutes of Health (NIH). Dr. Chou focused on fostering collaboration in research and practical efforts between the fields of public health and computational social science. She gave an overview of social media research at the NIH, discussing priority areas, detailing emerging opportunities for computer science, and offering a brief look at the funding portfolio. Finally, she outlined useful tips for a successful NIH-funded career.

Seven papers were presented, spanning a variety of health topics at the intersection of social groups or individuals and social media: assigning correct physician specialists in response to patient inquiries; promoting social ties and connectedness; analyzing the activities and outcomes of smoking cessation groups on Facebook; identifying risk markers for and

predicting the relapse of schizophrenia; comparing symptom reports for respiratory infection with Twitter timelines to determine how reliable an indicator of illness internet data can be; providing health-aware food recommendations; and studies in the integration of online and offline data in the sensitive problem domain of mental health. While the papers were diverse in their substantive topics, they shared common methodological challenges regarding external generalizability, feature selection, biases in data, and the transfer of findings between data source types. The workshop was also notable for bringing together researchers from three different countries.

Overall, there was lively discussion both during the talks and in the poster session, where all authors had the opportunity to present their work beyond the talk. This lively diversity is indicative of the growing social media and health community and the need for further events to build on the themes of this workshop.

Yelena Mejova, Rumi Chunara, Kyriaki Kalimeri, and Michael Paul served as cochairs of the workshop. This report was written by Yelena Mejova and Rumi Chunara.

## The Social Web for Environmental and Ecological Monitoring

The goal of the third international Social Web for Environmental and Ecological Monitoring workshop was to bring together researchers in academe and industry to discuss ideas, challenges, and solutions at the intersection of social media and environmental and ecological science.

The motivation behind the workshop was twofold. First, on the environmental science side, researchers need data about environmental and ecological processes, especially in the era of global climate change. While researchers continue to perform high-quality science with traditional sources of data like satellites and manual observations, these sources have limitations. Satellites can observe over large-scale spatial areas over long periods of time, but cannot observe ground-level, fine-grained phenomena like animal migration. Meanwhile, ground-based stations and manual observations collect fine-grained data about a wider range of phenomena, but are costly and restricted to small spatial areas.

Second, from the social web side, the exponential growth of social media has not only provided novel techniques for public communication and engagement, but has also produced unprecedented volumes of publicly available, user-generated content. These trends open new opportunities for ecological and environmental applications, both in terms of alternative data sources and of novel approaches for interacting with the public. Social media researchers are looking for new ways to apply the massive data



obtained from social media, as well as for novel communication and engagement techniques to help solve real-world problems. Ecological and environmental science could be one particularly striking application: user-generated media content provides an enormous amount of data that can be used in environmental models, while social media channels are an excellent way to engage the public.

But a challenge in fully realizing this application is the limited communication between the two fields. Most social media researchers have limited domain knowledge of environmental science, and thus struggle to identify important environmental-related problems and use cases. Meanwhile, ecological and environmental scientists may not be aware of the technical possibilities, limitations, and methods involved in social media analysis. The goal of this workshop was to help fill this gap by bringing together researchers interested in the intersection of the two fields, to begin to develop a research community around this exciting and emerging area. The workshop was the third in a series that started at ICWSM 2016 in Cologne, Germany, and continued at WebSci 2017 in Troy, New York.

In total, the workshop featured 10 talks covering a broad range of topics, including social media data mining and analysis, semantic content extraction, geolocated and network data, visual data mining, social networks, citizen science, and crowdsourcing. The speakers included Gretchen Lebuhn (San Francisco State University) with “The Great Sunflower Project”; Gerald Friedland (University of California, Berkeley) with “Field Studies with Multimedia Big Data: Opportunities and Challenges”; Shawn Newsam (University of California, Merced) with “Computer Science Meets the Environment: Three Projects”; Yu-Ru Lin (University of Pittsburgh) with “Event Analytics for Strengthening Community Resilience in a Cyber-Physical Society”; Qingkai Kong (University of California, Berkeley) with “MyShake: A Smartphone Seismic Network for Earthquake Early Warning and Beyond”; Felix Jan Hein Hol (Stanford University) with “Using Mobile Phones as Acoustic Sensors for High-Throughput Mosquito Surveillance”; Darian Frajberg (Politecnico di Milano) with “Augmented Reality and Artificial Intelligence for crowdsourced mountains monitoring”; Omar Alonso (Microsoft) with “Urban Maps of Social Activity”; Firoj Alam (Hamad Bin Khalifa University) with “CrisisMMD: Multimodal Twitter Datasets from Seven Natural Disasters”; and David Crandall (Indiana University) with “Tracking Natural Events through Social Media and Computer Vision.”

The workshop concluded with discussions and brainstorming on the main challenges of this research field and the prospective directions for future work. Participants agreed that social media analysis applied to environmental and ecological monitoring has been underexplored and that we

should work towards building a research community around this promising area.

David J. Crandall and Darian Frajberg organized the event and wrote this report.

## The ICWSM Science Slam

A *science slam* is an epic scientific event where scientists compete with short talks on their research. It’s just like a poetry slam, but with science instead of poems. Slammers were completely free to do whatever they wanted on stage. Everything was allowed, including slides, games — the more creative, the better! The only two rules were that the topic of the slam had to be related to social media and that the presentation could not take more than eight minutes. The 2018 ICWSM Science Slam was organized by Fabian Flöck and Cody Buntain.

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

**Rumi Chunara** is an assistant professor at New York University, USA.

**David J. Crandall** is an associate professor at Indiana University, Bloomington, USA.

**Darian Frajberg** is a PhD candidate at Politecnico di Milano, Italy.

**Megan French** is a graduate student in communication at Stanford University, USA.

**Bernard J. Jansen** is a principal scientist at the Qatar Computing Research Institute, Hamad Bin Khalifa University, Qatar.

**Juhi Kulshrestha** is a postdoctoral researcher at GESIS - Leibniz Institute for the Social Sciences, Cologne, Germany.

**Yelena Mejova** is a scientist at the Qatar Computing Research Institute, Hamad Bin Khalifa University, Qatar.

**Daniel M. Romero** is an assistant professor in the School of Information, University of Michigan, USA.

**Joni Salminen** is a postdoc at the Qatar Computing Research Institute, Hamad Bin Khalifa University, Qatar, and at the Turku School of Economics, Finland.

**Amit Sharma** is a researcher at Microsoft Research, India.

**Amit Sheth** is the executive director of Kno.e.sis Center at Wright State University, USA.

**Chenhao Tan** is an assistant professor in the Department of Computer Science, University of Colorado Boulder, USA.

**Samuel Hardman Taylor** is a graduate student in communication at Cornell University, USA.

**Sanjaya Wijeratne** is a PhD candidate at the Kno.e.sis Center, Wright State University, USA.