# Research Challenges of Digital Misinformation: Toward a Trustworthy Web

*Giovanni Luca Ciampaglia, Alexios Mantzarlis, Gregory Maus, and Filippo Menczer*

■ Abstract.

The deluge of online and offline misinformation is overloading the exchange of ideas upon which democracies depend. Fake news, conspiracy theories, and deceptive social bots proliferate, facilitating the manipulation of public opinion. Countering misinformation while protecting freedom of speech will require collaboration across industry, journalism, and academia. The Workshop on Digital Misinformation — held in May 2017 in conjunction with the International Conference on Web and Social Media in Montreal — was intended to foster these efforts. The meeting brought together more than 100 stakeholders from academia, media, and tech companies to discuss the research challenges implicit in building a trustworthy Web. Below we outline the main findings from the discussion.

■ Key Words

Misinformation, Fake News, Social Media, Trust, Credibility, Fact-Checking

## Introduction

The deluge of online misinformation is overloading the exchange of ideas upon which democracies depend. Many have argued that echo chambers are increasingly constricting the ability of alternative perspectives to provide a check on one's viewpoints (Sunstein 2009; Pariser 2011; Sunstein 2002).  Suffering fragmentation and declining public trust, the Fourth Estate struggles to carry out its traditional editorial role of distinguishing facts from fiction (Pew Research Center 2017). Within this context, fake news, conspiracy theories, and deceptive social bots proliferate, facilitating the manipulation of public opinion.

Countering misinformation while protecting freedom of speech will require collaboration between stakeholders across the tech industry, journalism, and academia. To foster such collaboration, the Workshop on Digital Misinformation was held in conjunction with the 2017 International Conference on Web and Social Media (ICWSM) in Montreal, on May 15, 2017. The meeting brought together more than 100 stakeholders from academia, media, and tech companies to discuss research challenges toward a trustworthy Web.
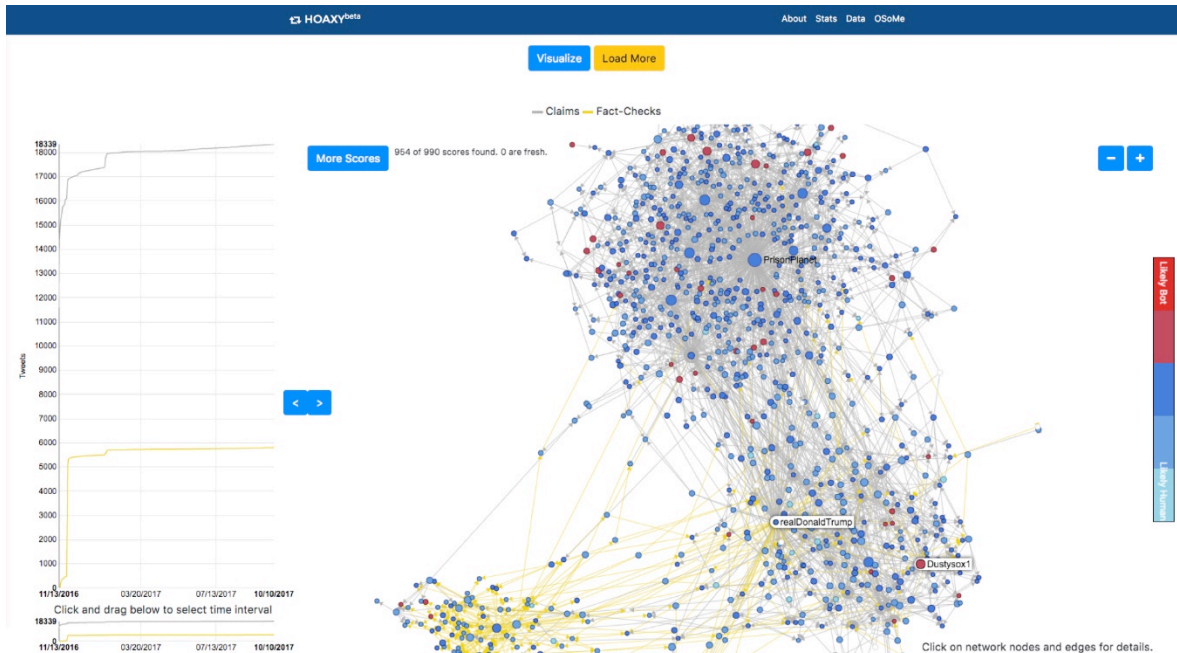
*Figure 1: Screenshot of the Hoaxy system (hoaxy.iuni.iu.edu). This is a visualization of the spread of fake news (purple) and fact checking (orange) articles related to the query "Three Million Votes Illegal Aliens" on Twitter. The left panel plots the number of tweets with links to these articles over time. The right panel show a diffusion network in which nodes represent Twitter accounts and edges represent the spread of the articles via retweets and mentions.*

The workshop opened with a showcase of tools for studying digital misinformation, developed by the Center for Complex Networks and Systems Research and the IU Network Science Institute. These are part of a growing suite of publicly available tools called the Observatory on Social Media (Davis et al. 2016). They include Hoaxy, a system for tracking competing claims and fact-checking that spread online (Shao et al. 2017) (Figure 1); Botometer, an AI system for detecting social bots on Twitter (Davis et al. 2016; Varol et al. 2017) (Figure 2); and unsupervised graph mining methods for automatically estimating factual accuracy from DBpedia (Shiralkar et al. 2017) (Figure 3). We then presented empirical data showing that, on social media, low-quality information often spreads more virally than high-quality information (Qiu et al. 2017). Factors that can explain this finding include the structural segregation and polarization in online social networks (Conover et al. 2011). The resulting 'echo chambers' are exacerbated by algorithms that personalize online experiences and hinder exposure to ideologically diverse sources of information (Nikolov et al. 2015). Other factors include information overload (Qiu et al. 2017), limited attention (Weng et al. 2012), popularity bias (Nematzadeh et al. 2017), and manipulation through social bots (Ferrara et al. 2016). One of the key question raised during discussion was how to empirically define "quality" of information in modern social media.

A lightning talk session was opened by BuzzFeed media editor Craig Silverman, who proposed a working definition of 'fake news' as "fabricated news intended to deceive with financial motive." Silverman emphasized the global scope of the issue, describing similar problems in Germany, Japan, Italy, and Myanmar, even if the social networks used to spread such news vary. Journalists are now aware of the scope of the problem and are reporting on fake news stories — though not always well. Silverman called for more collaboration between journalists and academic researchers for the study of misinformation; his team at Buzzfeed has developed a curated list of partisan news sites that is openly shared (Silverman 2017).

Political scientist and communication scholar Leticia Bode (Georgetown University) reported on a number of findings about correcting social media misinformation. Her work focuses on health and science communication, where the distinction between factual and opinionated claims is more clear compared to other domains such as political communication (Bode and Vraga 2015). Bode and her collaborators found that certain topics (e.g., GMOs) are easier to correct than others (e.g., vaccines and autism). She also found that 'social' fact-checking is more effective if it links to credible sources. Based on these findings, Bode recommended that news organizations should emphasize easily linked references and that corrections
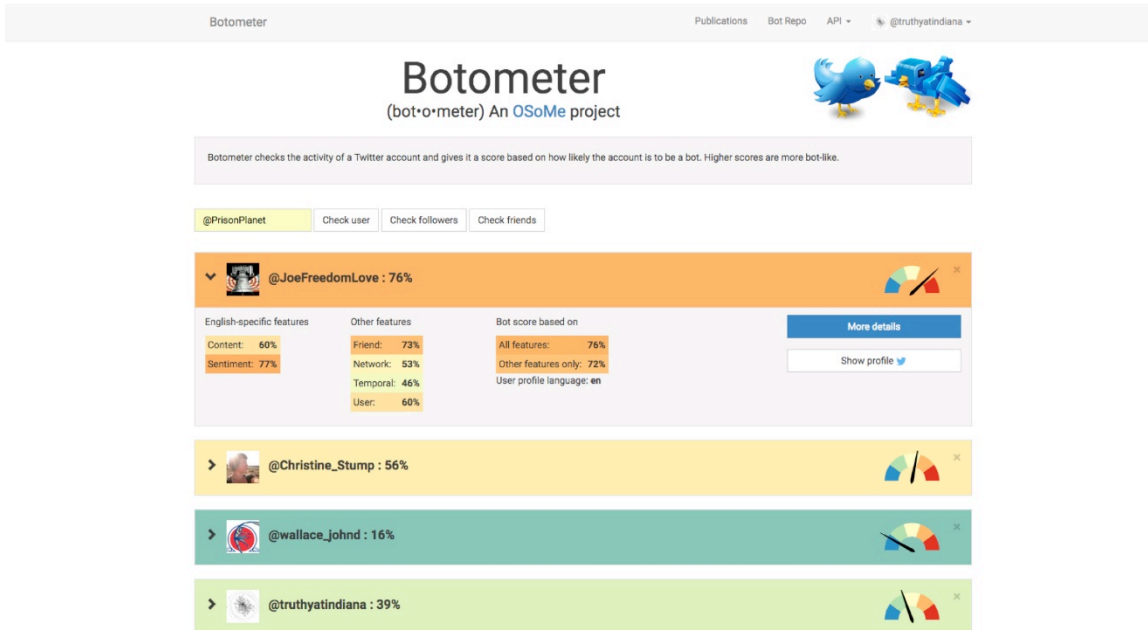
*Figure 2: Screenshot of the Botometer system (botometer.iuni.iu.edu). This is a supervised learning framework that calculates the likelihood that a given Twitter account is controlled by software (social bot).*

should be early and repeated. A new partnership model with social media platforms could satisfy these requirements, she concluded.

The perspective of a leading social media platform was given by Áine Kerr, leader of global journalism partnerships at Facebook. She started by sharing figures to appraise the scale at which Facebook's News Feed operates, with hundreds of million links shared weekly. Kerr noted that the quality of those links varies dramatically, and quoted Mark Zuckerberg's call for amplifying the good effects of social media and mitigating the bad. Facebook is pursuing this goal with four approaches: (1) disrupting the financial incentives for fake news; (2) developing new products to curb the spread of fake news, such as allowing users or third-party fact-checkers to flag posted stories as untrue or unverified; (3) helping people make informed decisions by educating them on how to spot fake news; and (4) launching the News Integrity Initiative, a partnership between industry and non-governmental organizations to promote media literacy (Mosseri 2016). Kerr noted that Facebook regularly engages with the research community via collaborative programs and grants, but acknowledged that there is a growing demand from third-party researchers for data to tackle the above problems; the organization is constantly refining its best practices around data sharing. A lively discussion followed about the commitment of platforms to curbing the spread of misinformation. For example, it was pointed out that more should be done to deal with abuses that exploit social bots, fake accounts, Facebook pages, and 'verified account' badges. It was also suggested that an API for access to public Facebook page data would be a great boon to the research community.

Computer Scientist Paul Resnick (University of Michigan) argued that factual corrections are often ineffective and slow; they rarely reach the people originally influenced by the misinformation. Moreover, we are exposed to information filtered by socio-technical mechanisms that largely prioritize popularity over accuracy, such as search engines, upvotes, and newsfeeds. To restore the balance in favor of accuracy, Resnick called for the development of reputation-based filtering mechanisms, and reported on his own work on gauging the trustworthiness of sources using a prediction market (Resnick and Sami 2008). In this scenario, third parties would verify the content shared by the sources, whose credibility would be automatically re-evaluated in light of new verifications. Thus, those who consistently share accurate information would have greater influence than those who do not. Resnick argued that a more widespread adoption of reputation-based filtering technologies would lead to more accurate content being shared.

Data scientist Rolf Fredheim (NATO StratCom Centre of Excellence) discussed the recent spate of misinformation operations in Europe (especially Russia-sponsored ones) and the problems created by the decreasing credibility of governments and the media in the eyes of the public, as well as the increasing use

of deceptive social bots (Fredheim 2017). He noted that while governments want a quick fix to the problem, comprehensive approaches are required, including adjusting the incentives for news production and consumption as well as providing education and awareness of misinformation. Fredheim insisted that social media need to be considered seriously as a tool of soft power and thus social media companies should be pressured to police their platforms.

## How to define and detect misinformation?

A first panel discussed the research challenge of defining and detecting misinformation in online settings. Computer scientist Jisun An (Qatar Computing Research Institute) demonstrated the economic threats of misinformation by discussing a 2013 incident in which hackers gained control of the Twitter feed of the Associated Press and used it to falsely announce an attack on President Obama, causing an immediate plunge in stock market prices. Computer scientist Meeyoung Cha (Korea Advanced Institute of Science and Technology) defined misinformation as information unverified at the time of circulation and later determined to be false, and proposed to employ AI algorithms for its detection (Zhao, Resnick, and Mei 2015; Ma et al. 2016). More theoretical work is necessary to understand why people spread rumors; this is less well understood than the incentives for creating rumors, particularly the financial ones.

Several statistical approaches for detecting misinformation require labeled instances to train machine learning models. Computer scientist James Caverlee (Texas A&M University) reported on the critical issue of producing such ground truth. He started by describing intuitive identification methods, such as the one according to which those users friending a fake account are themselves deserving of some scrutiny. He argued that while powerful, these simple heuristics can only address a fraction of the issue. When it comes to more complex claims, fact-checkers are key to identifying the ground truth. However, they cannot cope with the sheer volume and variety of misinformation. Caverlee proposed exploiting aggregated signals to infer the reliability of a given piece of content, such as the reply-to-retweet ratio to flag controversy (Alfifi and Caverlee 2017; Kaghazgaran, Caverlee, and Alfifi 2017). Caverlee warned that crowdsourced information can be easily manipulated; for example, it is easy to recruit workers for astroturfing. He noted that the Chinese government often fabricates social media posts in an attempt to prevent the public from discussing civic issues. Caverlee called for more research on the problem of identifying the intent behind social media posts. Progress in this area could lead to tools for distinguishing organic conversations from covert coordinated campaigns on social media (Varol et al. 2017).

Computer scientist Qiaozhu Mei (University of Michigan) highlighted how hacked accounts can bypass reputation systems, challenging solutions suggested by previous speakers. He noted that while spotting known rumors is easy after the fact, identifying (false) rumors in their early stages is difficult. He proposed a four-part approach: (1) identifying emerging rumors early by mining search engine logs for questions about a statement's truthfulness; (2) retrieving all the posts about a rumor; (3) analyzing the impact of a rumor and its correction via visualization; and (4) predicting its future spread through deep learning (Zhao, Resnick, and Mei 2015; Li et al. 2017; Qazvinian et al. 2011; Resnick et al. 2014).

Computer scientist Eni Mustafaraj (Wellesley College) discussed an early example of fake news tactics, namely, a Twitter "bomb" against Massachusetts senate candidate Martha Coakley in 2010 (Mustafaraj and Metaxas 2010). A group of Twitter bots used hashtags targeted towards specific communities in a coordinated campaign to spread negative information about the candidate. Mustafaraj compared this attack with recent ones, leveraging fake Facebook accounts to target specific Facebook groups and spread links to fake news stories (Mustafaraj and Metaxas 2017). She examined three motives for spreading such inaccurate information: financial, political, and ideological/cultural (including prejudices like sexism and xenophobia). She proposed that social media platforms should highlight provenance information about sources to help users determine their intents and trustworthiness. Finally, Mustafaraj urged platforms to provide researchers with data about how recipients of misinformation engage with it.

Information scientist Kate Starbird (University of Washington) delineated different types of rumor in a crisis situation, such as hoaxes, collective sense-making, and conspiracy theories (Arif et al. 2016). She categorized methods for detecting misinformation based on linguistic, network, account, URL domain, and crowdsourced features (Maddock et al. 2015; Huang et al. 2015; Dailey and Starbird 2014).

Workshop participants responded to an online Wiki survey and identified five major research challenges related to the question of how to define and detect misinformation (Table 1).

# Survey of Participant Priorities by Panel

**Panel 1: Defining and Detecting Misinformation**

| |
|---|
| 1. Identifying and promoting reliable information instead of focusing on disinformation |
| 2. Tracking variants of debunked claims |
| 3. Developing reputation scores for publishers |
| 4. Creating an automated trustmark to promote journalistic integrity |
| 5. Collecting reliable crowdsourcing signals |

**Panel 2: Cognitive, Social, and Technological Biases Creating Vulnerabilities to Misinformation**

| |
|---|
| 1. Investigating the use of language, images, and design in misinformation persuasiveness |
| 2. Validating model predictions via field experiments |
| 3. Studying the roles of algorithmic mechanisms in the spread of misinformation |
| 4. Translating research findings into policy recommendations |
| 5. Accessing behavioral data from social media platforms |

**Panel 3: Countermeasure Feasibility, Effectiveness, and Responsibility**

| |
|---|
| 1. Support and scaffold critical thinking |
| 2. Increase prominence and availability of fact-checking information |
| 3. Design trust & reputation standards for news sources and social media users |
| 4. Build tools to track the provenance of digital content |
| 5. Develop computational tools to support fact-checking |

*Table 1: After each panel, participants in the workshop proposed and chose top priorities for each topic via a dynamic online head-to-head ranking system. This table records the top five results for each.*

## How to best study the cognitive, social, and technological biases that make us vulnerable to misinformation?

A second panel addressed research challenges in studying cognitive, social, and technological biases that create vulnerabilities to misinformation. Stanford University historian Joel Breakstone cited a study showing that students have difficulty distinguishing the reliability of news stories (McGrew et al. 2016). In this study, students' assessments of the veracity of the content they were presented were found to rely primarily on the appearance of the article rather than on consideration of the source. The same study also found that students were unable to distinguish native advertising from real news stories 80% of the time. Breakstone concluded that the fight against misinformation is one in which we all must take part, not just big tech companies.

Communications scholar R. Kelly Garrett (Ohio State University) sought to clarify several points in the discussion. First, he distinguished between holding a belief and being ignorant of the evidence against it, citing statistics on the number of people who know the scientific consensus on global warming but reject it (Funk and Kennedy 2016). However, he noted that online partisan news lead people to reject evidence, due to these outlets' emotional pull (Garrett, Weeks, and Neo 2016). While social media increase the profile of misinformation, it remains unclear how much this actually shifts public opinions.

Physicist Kristina Lerman (University of Southern California) emphasized that humans have limited information processing capabilities, which makes it impossible to keep up with the growing volume of information, resulting in reliance on simple cognitive heuristics. These heuristics may in turn amplify certain cognitive biases. Lerman reported on her studies about popularity, engagement, and position bias (Kang and Lerman 2015; Lerman and Hogg 2014). The latter is the idea that people pay more attention to what is at the top of a list or screen. Experimental trials show that news stories at the top of a list are 4-5 times more likely to be shared. Lerman also discussed how social media reinforce network biases, creating echo chambers that distort our perceptions.

Economist David Rothschild (Microsoft Research) proposed to address misinformation as a market problem, framing it in terms of outcomes such as exposure to information and impact on opinion formation and decision making. Rothschild noted that research in the field tends to focus on what content people consume rather than the more difficult matter of how they actually absorb that information. He also questioned

whether mass ignorance on a particular issue may be more harmful than the consumption of fake news about that issue. Research may also be distorted by over-reliance on Twitter data, Rothschild suggested, which may not be as representative of news consumption by the general population as Facebook or television (Diaz et al. 2016).

Computer scientist Kazutoshi Sasahara (Nagoya University) presented work in progress on a simple model of online social network dynamics. The model demonstrates that online echo chambers are inevitable within the current social media mechanisms for content sharing, which tend to clusters individuals into segregated and polarized groups.

Workshop participants responded to an online Wiki survey and identified five major research challenges related to the question of how to best study the cognitive, social, and technological biases that make us vulnerable to misinformation (Table 1).

## What countermeasures are most feasible and effective and who can best deliver them?

The final panel discussed countermeasures against misinformation, as well as who can best deliver them. Computational journalist Nick Diakopoulos (University of Maryland) identified three relevant actors to be considered: tech platforms, individuals, and civil society. He discussed which combinations of the three groups could be most effective in combating fake news. Platforms are particularly powerful, but they raise the concern of influencing public discourse through their algorithms — although this could be mitigated through algorithmic transparency (Diakopoulos 2017; Diakopoulos and Koliska 2017). Civil society and individuals alone cannot fact-check everything. Diakopoulos concluded that the best partnership would be between civil society and platforms.

Fact-checker David Mikkelson (Snopes.com) argued that fake news is only as problematic as poor journalism: the very news outlets that are supposed to question and disprove misinformation often help spread it (Mikkelson 2016).

Computer scientist Tim Weninger (University of Notre Dame) addressed the lack of research on Reddit, a far larger platform than Twitter. He reported on findings that initial votes on posts have a strong impact on their final visibility, allowing coordinated attacks to game the system through a snowballing effect (Glenski and Weninger 2017). Weninger also found that a large portion of Reddit users merely scan headlines; most up or down votes are cast without even viewing the content (Glenski, Pennycuff, and Weninger 2017).

Computer scientist Cong Yu (Google Research) described the use of semantic web annotations such as the Schema.org *ClaimReview* markup to help surface fact checks of popular claims on search engines (Kosslyn and Yu 2017). He argued that artificial intelligence can be a powerful tool to promote quality and trust in information. However, Yu recognized that users play a role in the spread of misinformation, which may be the most challenging problem to address.

Communication scientist Melissa Zimdars (Merrimack College) reported on efforts to collectively categorize news sources in an open-source fashion (opensources.co). She also recounted how, ironically, her research became the target of a fake news campaign.

Workshop participants responded to an online Wiki survey and identified five major research challenges related to the question of what countermeasures are most feasible/effective and who can best deliver them (Table 1).

## Conclusions

Unfortunately, Artificial Intelligence is increasingly being exploited to manipulate public opinion. For example, sophisticated social bots can autonomously interact with social media users in an attempt to influence them or expose them to misinformation. Advances in machine generation of realistic video and voice have been already identified as the likely next-generation weapons in the digital misinformation arsenal (Suwajanakorn, Seitz, and Kemelmacher-Shlizerman 2017; Thies et al. 2016).

The good news is that AI can also play an important role in defending us from attacks against the integrity of the information space. In such an arms race, advances in supervised and unsupervised machine learning, representation learning, and natural language processing will be needed to help meet the above challenges.
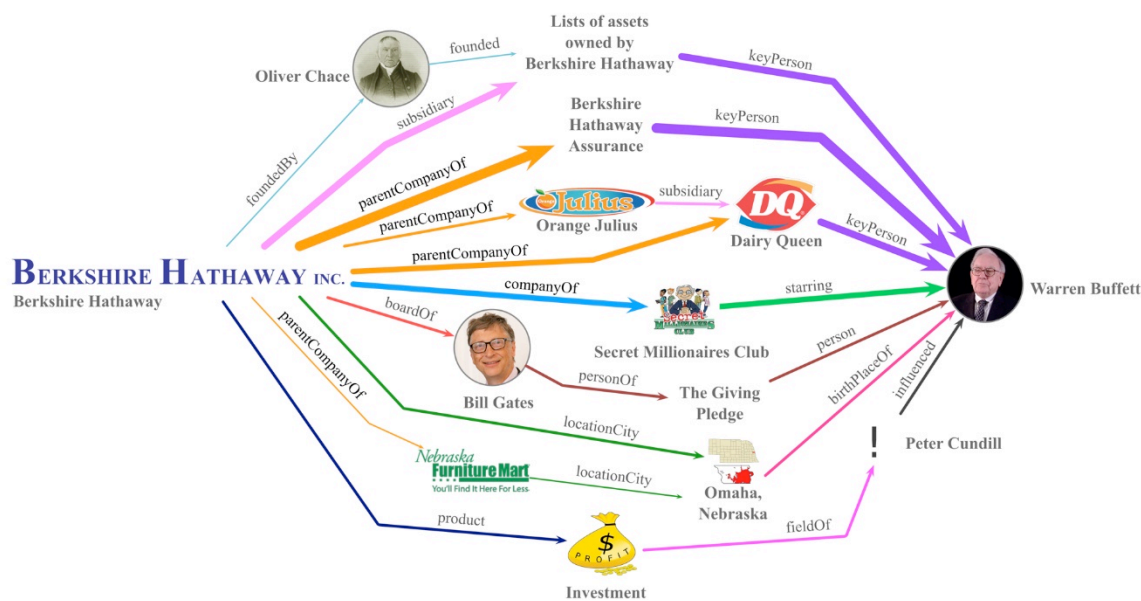
*Figure 3: An example of connections identified by a computational fact-checking algorithm to verify a claim about a company and its CEO. A flow algorithm is used to identify relational paths in a knowledge network. The width of an edge is proportional to the flow.*

Another area where more AI research is needed is the study of algorithmic bias. Social media platforms employ sophisticated ranking, filtering, and recommendation methods that are increasingly powered by cutting-edge AI algorithms. Unfortunately, these algorithms are also vulnerable to manipulation due to their focus on engagement and popularity, leading to echo chambers and selective exposure that amplify our own cognitive and social biases. A significant challenge will be to improve algorithms to take into account signals of trustworthiness and reliability.

Finally, reporters and fact-checking organizations are in great need of tools to help them manage the volume of digital misinformation at scale. Semantic Web technologies are reaching maturity and can be leveraged toward the development of computational fact-checking tools. Automatic extraction of semantic data from text provides large knowledge bases which, in combination with inference and network mining techniques, could accelerate the tasks of verification and news gathering, for example by connecting claims with relevant contextual information and previous analyses (Figure 3). Special emphasis should be placed upon interpretable results, that can aid both journalists and the general public make sense of the information to which they are exposed.

The ultimate goal of the workshop was to bootstrap a long-lasting initiative between various sectors (namely industry, academia, journalism, and civil society) with the aim of building a trustworthy Web. Follow-up discussion and further collaborative activity is currently under way through online community spaces. We believe that support from both private foundations and federal agencies will be a key ingredient for the success of future collaborative activities, the scope of which must include research, education, and policymaking.

References

Alfifi, Majid, and James Caverlee. 2017. "Badly Evolved? Exploring Long-Surviving Suspicious Users on Twitter." *Proceedings of SocInfo 2017, International Conference* on *Social Informatics. Lecture Notes in Computer Science,* vol. 10539, pp. 218–233. Available at https://doi.org/10.1007/978-3-319-67217-5_14

Arif, Ahmer, Kelley Shanahan, Fang-Ju Chou, Yoanna Dosouto, Kate Starbird, and Emma Spiro. 2016. "How Information Snowballs: Exploring the Role of Exposure in Online Rumor Propagation." *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*, 466–477. Available at https://doi.org/10.1145/2818048.2819964

Bode, Leticia, and Emily K. Vraga. 2015. "In Related News, That Was Wrong: The Correction of Misinformation Through Related Stories Functionality in Social Media." *Journal of Communication* 65 (4): 619–638. Available at https://doi.org/10.1111/jcom.12166

Conover, M, J Ratkiewicz, M Francisco, B Gonçalves, F Menczer, and A Flammini. 2011. "Political Polarization on Twitter." In *Proc. Fifth International AAAI Conference on Weblogs and Social Media (ICWSM)*, 89–96. Available at https://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2847

Dailey, Dharma, and Kate Starbird. 2014. "Journalists as Crowdsourcerers: Responding to Crisis by Reporting with a Crowd." *Computer Supported Cooperative Work* 23 (4–6): 445–481. Available at https://doi.org/10.1007/s10606-014-9208-z

Davis, Clayton A., Giovanni Luca Ciampaglia, Luca Maria Aiello, Keychul Chung, Michael D. Conover, Emilio Ferrara, Alessandro Flammini, et al. 2016. "OSoMe: The IUNI Observatory on Social Media." *PeerJ Computer Science* 2:e87. Available at https://doi.org/10.7717/peerj-cs.87

Davis, Clayton Allen, Onur Varol, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer. 2016. "BotOrNot: A system to evaluate social bots." In *Proceedings of the 25th International Conference Companion on World Wide Web (WWW'16) Companion*, 273–274. Available at https://doi.org/10.1145/2872518.2889302

Diakopoulos, Nicholas. 2017. "Enabling Accountability of Algorithmic Media: Transparency as a Constructive and Critical Lens," In: Cerquitelli T., Quercia D., Pasquale F. (eds) *Transparent Data Mining for Big and Small Data. Studies in Big Data,* vol 32, pp. 25–43, Springer. Available at https://doi.org/10.1007/978-3-319-54024-5_2

Diakopoulos, Nicholas, and Michael Koliska. 2017. "Algorithmic Transparency in the News Media." *Digital Journalism* 5 (7): 809–828. Available at https://doi.org/10.1080/21670811.2016.1208053

Diaz, Fernando, Michael Gamon, Jake M. Hofman, Emre Kıcıman, and David Rothschild. 2016. "Online and Social Media Data As an Imperfect Continuous Panel Survey." *PLoS ONE* 11 (1): e0145406. Available at https://doi.org/10.1371/journal.pone.0145406

Ferrara, Emilio, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. 2016. "The Rise of Social Bots." *Communications of the ACM* 59 (7): 96–104. Available at https://doi.org/10.1145/2818717

Fredheim, Rolf. 2017. "Robotrolling." NATO Strategic Communications Centre of Excellence. Available at https://www.stratcomcoe.org/robotrolling-20171

Funk, Carey, and Brian Kennedy. 2016. "The Politics of Climate." *Pew Research Center*. Available at http://www.pewinternet.org/2016/10/04/the-politics-of-climate/

Garrett, R. Kelly, Brian E. Weeks, and Rachel L. Neo. 2016. "Driving a Wedge Between Evidence and Beliefs: How Online Ideological News Exposure Promotes Political Misperceptions." *Journal of Computer-Mediated Communication* 21 (5): 331–348. https://doi.org/10.1111/jcc4.12164

Glenski, Maria, Corey Pennycuff, and Tim Weninger. 2017. "Consumers and Curators: Browsing and Voting Patterns on Reddit." *IEEE Transactions on Computational Social Systems*, in press. Available at https://doi.org/10.1109/TCSS.2017.2742242

Glenski, Maria, and Tim Weninger. 2017. "Rating Effects on Social News Posts and Comments." *ACM Transactions on Intelligent Systems and Technology* 8 (6): 78. Available at https://doi.org/10.1145/2963104

Huang, Y Linlin, Kate Starbird, Mania Orand, Stephanie A Stanek, and Heather T Pedersen. 2015. "Connected through Crisis: Emotional Proximity and the Spread of Misinformation Online." *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*, 969–p80. Available at https://doi.org/10.1145/2675133.2675202

Kaghazgaran, P., J. Caverlee, and M. Alfifi. 2017. "Behavioral Analysis of Review Fraud: Linking Malicious Crowdsourcing to Amazon and beyond." *Proceedings of the 11th International AAAI Conference on Web and Social Media (ICWSM 2017)*, 560–563. Available at https://www.aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/view/15659

Kang, Jeon Hyung, and Kristina Lerman. 2015. "VIP: Incorporating Human Cognitive Biases in a Probabilistic Model of Retweeting." *Proc. International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction (SBP 2015). Lecture Notes in Computer Science* 9021: 101–110. Available at https://doi.org/10.1007/978-3-319-16268-3_11

Kosslyn, Justin, and Cong Yu. 2017. "Fact Check Now Available in Google Search and News around the World." *Google Blog*. Available at https://www.blog.google/products/search/fact-check-now-available-google-search-and-news-around-world/

Lerman, Kristina, and Tad Hogg. 2014. "Leveraging Position Bias to Improve Peer Recommendation." *PLoS ONE* 9 (6): e98914. Available at https://doi.org/10.1371/journal.pone.0098914

Li, Cheng, Jiaqi Ma, Xiaoxiao Guo, and Qiaozhu Mei. 2017. "DeepCas: An End-to-End Predictor of Information Cascades." In *Proceedings of the 26th International Conference on World Wide Web (WWW '17)*, 577–586. Available at https://doi.org/10.1145/3038912.3052643

Ma, Jing, Wei Gao, Prasenjit Mitra, Sejeong Kwon, Bernard J Jansen, Kam-Fai Wong, and Meeyoung Cha. 2016. "Detecting Rumors from Microblogs with Recurrent Neural Networks." *Proc. 25th International Joint Conference on Artificial Intelligence (IJCAI'16)*, 3818–3824. Available at http://www.ijcai.org/Abstract/16/537

Maddock, Jim, Kate Starbird, Haneen J. Al-Hassani, Daniel E. Sandoval, Mania Orand, and Robert M. Mason. 2015. "Characterizing Online Rumoring Behavior Using Multi-Dimensional Signatures." *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*, 228–241. Available at https://doi.org/10.1145/2675133.2675280

McGrew, Sarah, Teresa Ortega, Joel Breakstone and Sam Wineburg. 2017. "The Challenge That's Bigger Than Fake News." *American Educator* (Fall): 4–9.

Mikkelson, David. 2016. "We Have a Bad News Problem, Not a Fake News Problem." Snopes. 2016. Available at https://www.snopes.com/2016/11/17/we-have-a-bad-news-problem-not-a-fake-news-problem/

Mosseri, Adam. 2016. "News Feed FYI: Addressing Hoaxes and Fake News." Facebook. Available at https://newsroom.fb.com/news/2016/12/news-feed-fyi-addressing-hoaxes-and-fake-news/

Mustafaraj, Eni, and Panagiotis Takis Metaxas. 2010. "From Obscurity to Prominence in Minutes: Political Speech and Real-Time Search." *Proc. WebSci10: Extending the Frontiers of Society On-Line*. Available at http://repository.wellesley.edu/computersciencefaculty/9/

Mustafaraj, Eni, and Panagiotis Takis Metaxas. 2017. "The Fake News Spreading Plague: Was It Preventable?" *Proceedings of the ACM Conference on Web Science (WebSci'17)*, pp. 235–239. Available at https://doi.org/10.1145/3091478.3091523

Nematzadeh, Azadeh, Giovanni Luca Ciampaglia, Filippo Menczer, and Alessandro Flammini. 2017. "How Algorithmic Popularity Bias Hinders or Promotes Quality." Preprint arXiv:1707.00574. Available at https://arxiv.org/abs/1707.00574

Nikolov, Dimitar, Diego F. M. Oliveira, Alessandro Flammini, and Filippo Menczer. 2015. "Measuring Online Social Bubbles." *PeerJ Computer Science* 1: e38. Available at https://doi.org/10.7717/peerj-cs.38

Pariser, E. 2011. *The Filter Bubble: How the New Personalized Web Is Changing Wha We Read and How We Think*. Penguin.

Pew Research Center. 2017. "State of the News Media." Available at http://www.pewresearch.org/topics/state-of-the-news-media/

Qazvinian, Vahed, Emily Rosengren, Dragomir R Radev, and Qiaozhu Mei. 2011. "Rumor Has It:

Identifying Misinformation in Microblogs." *Proc. Conference on Empirical Methods in Natural Language Processing (EMNLP'11)*, 1589–1599. Available at https://aclanthology.info/papers/D11-1147/d11-1147

Qiu, Xiaoyan, Diego F. M. Oliveira, Alireza Sahami Shirazi, Alessandro Flammini, and Filippo Menczer. 2017. "Limited Individual Attention and Online Virality of Low-Quality Information." *Nature Human Behaviour* 1: 0132. Available at http://rdcu.be/tIa2

Resnick, Paul, Samuel Carton, Souneil Park, Yuncheng Shen, and Nicole Zeffer. 2014. "RumorLens : A System for Analyzing the Impact of Rumors and Corrections in Social Media." *Computation + Journalism Symposium*. Available at http://nicole.zeffer.com/cj2014.pdf

Resnick, Paul, and Rahul Sami. 2008. "Manipulation-Resistant Recommender Systems through Influence Limits." *ACM SIGecom Exchanges* 7 (3): 10. Available at https://doi.org/10.1145/1486877.1486887

Shao, Chengcheng, Giovanni Luca Ciampaglia, Onur Varol, Alessandro Flammini, and Filippo Menczer. 2017. "The Spread of Fake News by Social Bots." Preprint arXiv:1707.07592. Available at https://arxiv.org/abs/1707.07592

Shiralkar, Prashant, Alessandro Flammini, Filippo Menczer, and Giovanni Luca Ciampaglia. 2017. "Finding Streams in Knowledge Graphs to Support Fact Checking." *Proc. IEEE International Conference on Data Mining (ICDM)*. Available at https://arxiv.org/abs/1708.07239

Silverman, Craig. 2017. "Inside The Partisan Fight For Your News Feed." Available at https://github.com/BuzzFeedNews/2017-08-partisan-sites-and-facebook-pages

Sunstein, C.R. 2002. "The Law of Group Polarization." *The Journal of Political Philosophy* 10 (2): 175–195. Available at https://doi.org/10.1111/1467-9760.00148

———. 2009. *Republic.com 2.0*. Princeton University Press.

Suwajanakorn, Supasorn, Steven M. Seitz, and Ira Kemelmacher-Shlizerman. 2017. "Synthesizing Obama: Learning Lip Sync from Audio." *ACM Trans. Graphics* 36 (4): 95. Available at http://dx.doi.org/10.1145/3072959.3073640

Thies, Justus, Michael Zollhöfer, Marc Stamminger, Christian Theobalt, and Matthias Nießner. 2016. "Face2Face: Real-time Face Capture and Reenactment of RGB Videos." *Proc. Computer Vision and Pattern Recognition (CVPR)*, IEEE. Available at http://www.niessnerlab.org/papers/2016/1facetoface/thies2016face.pdf

Varol, Onur, Emilio Ferrara, Clayton A. Davis, Filippo Menczer, and Alessandro Flammini. 2017. "Online Human-Bot Interactions: Detection, Estimation, and Characterization." *Proceedings of the Eleventh International AAAI Conference on Web and Social Media (ICWSM)*. Available at https://aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/view/15587

Varol, Onur, Emilio Ferrara, Filippo Menczer, and Alessandro Flammini. 2017. "Early Detection of Promoted Campaigns on Social Media." *EPJ Data Science* 6: 13. Available at https://doi.org/10.1140/epjds/s13688-017-0111-y

Weng, L., A. Flammini, A. Vespignani, and F. Menczer. 2012. "Competition among Memes in a World with Limited Attention." *Scientific Reports* 2: 335. Available at https://doi.org/10.1038/srep00335

Zhao, Zhe, Paul Resnick, and Qiaozhu Mei. 2015. "Enquiring Minds: Early Detection of Rumors in Social Media from Enquiry Posts." *Proceedings of the 24th International Conference on World Wide Web (WWW '15)*, 1395–1405. Available at https://doi.org/10.1145/2736277.2741637

**Giovanni Luca Ciampaglia** is an assistant research scientist at the Indiana University Network Science Institute (IUNI). Before joining Indiana University, he was an analyst for the Wikimedia Foundation, and a research associate at the Swiss Federal Institute of Technology in Zurich. He is interested in information quality in cyber-human systems, in particular trustworthiness and reliability of information in intelligent systems. His work has been covered in major news outlets, including the *Wall Street Journal*, the *Economist*, *Wired*, *MIT Technology Review*, *NPR*, and *CBS News*.

**Alexios Mantzarlis** leads the International Fact-Checking Network at the Poynter Institute. Alexios has helped draft the fact-checkers' code of principles, shepherded a partnership between third-party fact-checkers and Facebook, testified to the Italian Chamber of Deputies on the "fake news" phenomenon, and helped launch International Fact-Checking Day. He previously served as Managing Editor of Pagella Politica and FactCheckEU, respectively Italy's main political fact-checking website and the EU's first multilingual crowd-checking project. While at Pagella Politica, he presented weekly fact-checking segments on Virus, a prime time talk show airing on Italy's national broadcaster RAI 2. Before becoming a fact-checker, he worked for the United Nations and the Italian Institute for International Political Studies.

**Gregory Maus** is a PhD student in Informatics at Indiana University, Bloomington. His work on social bots, personal data brokers, and algorithmic sociological influence has been featured in *IEEE Spectrum*, *Foreign Affairs*, and *Real Clear Technology*.

**Filippo Menczer** is a professor of Informatics and Computer Science at Indiana University, Bloomington. He holds a *Laurea* in Physics from the Sapienza University of Rome and a Ph.D. in Computer Science and Cognitive Science from the University of California, San Diego. Dr. Menczer is an ACM Distinguished Scientist. He previously served as division chair and director of the Center for Complex Networks and Systems Research at the IUB School of Informatics and Computing. His research focuses on Web and data science, social network analysis, social computation, Web mining, and modeling of complex information networks. His work on the spread of information and misinformation in social media has been covered in many US and international news sources, including *The New York Times*, *Wall Street Journal*, *Washington Post*, *NPR*, *PBS*, *CNN*, *BBC*, *Economist*, *Guardian*, *Atlantic*, *Reuters*, *Science*, and *Nature*. Menczer serves as associate editor of the *Network Science* journal and on the editorial boards of *EPJ Data Science* and *PeerJ Computer Science*.