

Overview of ROMCIR 2021: Workshop on Reducing Online Misinformation through Credible Information Retrieval

Fabio Saracco^a, Marco Viviani^b

^aIMT School for Advanced Studies, Piazza S. Ponziano, 6 – 55100 Lucca, Italy

^bUniversity of Milano-Bicocca (DISCo – IKR3 Lab), Edificio U14, Viale Sarca, 336 – 20126, Milan, Italy

Abstract

The 2021 Workshop on Reducing Online Misinformation through Credible Information Retrieval (ROMCIR 2021), as part of the satellite events of the 43rd European Conference on Information Retrieval (ECIR 2021), is concerned with providing users with access to genuine information, to mitigate the information disorder phenomenon characterizing the current online digital ecosystem. This problem is very broad, as it concerns different information objects (e.g., Web pages, online accounts, social media posts, etc.) on different platforms, and different domains and purposes (e.g., detecting fake news, retrieving credible health-related information, reducing propaganda and hate-speech, etc.). In this context, all those approaches that can serve, from different perspectives, to tackle the credible information access problem, find their place. Hence, this overview of the Workshop describes its motivations, scientific objectives, topics of interest, accepted papers, organization and organizational team.

Keywords

Information Disorder, Credibility, Information Retrieval

1. Motivations

Nowadays, we are more and more aware of the problems that can arise from coming into contact with different kinds of misleading contents that are propagated online, especially through social media platforms [1, 2]. False news can, for example, influence public opinion in political and financial choices [3]; false reviews can promote substandard products or, on the contrary, destroy florid economic activities by means on discredit campaigns [4]; unverified medical information can lead people to follow behaviors that can be harmful to their own health and to that of society as a whole (let us think, for example, of the risk of following negationism hypotheses in the context of the recent COVID-19 pandemic) [5].


This scenario is due to the so-called *information disorder* phenomenon [6], which indicates the proliferation of different forms of (online) communication pollution, encompassing *dis-*, *mis-*, and *mal-*information. Specifically, misinformation is the spread of false content resulting


ROMCIR 2021: Workshop on Reducing Online Misinformation through Credible Information Retrieval, held as part of ECIR 2021: the 43rd European Conference on Information Retrieval

✉ fabio.saracco@imtlucca.it (F. Saracco); marco.viviani@unimib.it (M. Viviani)

🌐 <https://www.imtlucca.it/en/fabio.saracco> (F. Saracco); <http://www.ir.disco.unimib.it/people/marco-viviani/> (M. Viviani)

🆔 0000-0003-0812-5927 (F. Saracco); 0000-0002-2274-9050 (M. Viviani)

 © 2021 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

from the spreader's ignorance; disinformation is a form of intentional sharing of false content to produce harm; malinformation indicates the spread of (private) information that is based on reality, having the same harmful intent (e.g., the despicable act of revenge porn).

Access to this non-genuine information is made easier and easier to the fact that, from a technological point of view, information is produced at a speed and volume never seen before, almost without any trusted traditional intermediary [7, 8]. Faced with this huge amount of information, and the uncertainty associated with its degree of veracity, human cognitive abilities are not always sufficient to take well-informed decisions [9].

In this context, it is clear that the problem of guaranteeing access to genuine information online needs to find effective solutions, despite (and precisely because of) it is very broad, as it concerns different information objects (e.g., Web pages, online accounts, social media posts, etc.), different online platforms (e.g., Web portals, social networking services, question-answering systems, etc.), and different domains and purposes (e.g., detecting fake news, retrieving credible health-related information, reducing propaganda and hate-speech, etc.).

2. Scientific Objectives

The key goal of the Workshop is to encourage a discussion between researchers also belonging to different disciplines, and propose innovative solutions, about the problem of guarantee to users access to credible information that does not distort their perception of reality. In recent years, despite numerous approaches have been proposed to tackle the considered issue in different contexts, and for different purposes, we are still a long way from having found completely effective and domain-independent solutions.

The problem is still of great interest with respect to many open issues, such as the access to and retrieval of credible information, the early detection of dis-/mis-/mal-information, the development of solutions that can be understood by final users (explainable AI), the study of the problem in the health-related field, the relationship between security, privacy and credibility in information access and dissemination.

In this scenario, the role of researchers working in the fields of Information Retrieval, Social Computing, Social Sciences, and other related research areas, is crucial to investigate such open issues, providing users with automatic but understandable tools to help them come into contact with genuine information.

3. Topics of Interest

All those approaches that can serve, from different perspectives, to tackle the credible information access problem, find their place in ROMCIR. Specifically, the topics of interest include, but are not limited to:

- Access to/retrieval of credible information;
- Bias evaluation and detection;
- Bot/spam/troll detection;

- Computational fact-checking;
- Crowdsourcing for information credibility assessment;
- Deep fake analysis and detection;
- Disinformation/misinformation/malinformation analysis and detection;
- Evaluation strategies to assess information credibility;
- Fake news/fake reviews detection, propaganda identification and analysis;
- Filter bubbles, echo chambers, and information polarization online;
- Harassment/bullying/hate-speech detection;
- Security, privacy, and information credibility;
- Sentiment/emotional analysis, and stance detection;
- Trust, reputation, and information credibility;
- Understanding and guiding the societal reaction in the presence of dis-/mis-/mal-information.

Both theoretical studies, model-driven, and data-driven approaches, supported by publicly available datasets, are more than welcome.

4. Submissions

The ROMCIR 2021 Workshop received 15 submissions, of which 6 were accepted, so with an acceptance rate of 40%. The accepted articles, collected in these Proceedings, consider different problems. There are issues tangentially related to credibility and Information Retrieval, such as those of authorship verification and bias detection in science evaluation. Furthermore, the problems of opinion mining and misinformation identification are tackled, such as those of hate speech detection and claim verification. Finally, the problem of the access to credible information is considered, by proposing the definition of systems to support users in retrieving genuine news, and the study of new IR methods able to consider the credibility of the data collected in the retrieval process.

In particular, the article by Zhang et al., entitled: *Improving Authorship Verification using Linguistic Divergence*, proposes an unsupervised solution to the authorship verification task that utilizes pre-trained deep language models to compute a new metric called DV-Distance. The proposed metric is a measure of the difference between two authors that takes into account the knowledge transferred from the pre-trained model. Bethencourt et al., in: *Bias and truth in science evaluation: a simulation model of grant review panel discussions*, use social simulation methods to study the discussion dynamics in peer review panels that seek to find the true merit of each submission. Using a simulation model, they explore whether a combination of assimilative influence and various kinds of biases could reproduce the opinion changes observed in real-world peer-review panel discussions.

In the article: *Consumption of Hate Speech on Twitter: A Topical Approach to Capture Networks of Hateful Users*, Gupta et al. attempt to track the dissemination of hate speech on Twitter. The authors argue that hate is not a blanket category, but exists across multiple topics. Hence, they use topic modelling to unearth the latent topics in tweets and an ensemble classification model to capture various nuances of hate speech. Hatua et al., in their article: *On the Feasibility of Using GANs for Claim Verification – Experiments and Analysis*, explore fact checking and claim verification by employing a Generative Adversarial Network (GAN) based model on the Fact Extraction and VERification (FEVER) dataset [10].

Supporting verification of news articles with automated search for semantically similar articles, by Gupta et al., describes an evidence retrieval approach to handle fake content. A system to find semantically similar articles to a given news article from selected trusted sources is developed; this way, users are helped in finding supporting evidence and, thus, manual search work can be reduced. Finally, Denaux and Gomez-Perez in: *Sharing Retrieved Information using Linked Credibility Reviews* summarise an existing work that provides a conceptualisation and exchange format for representing the credibility of retrieved verified information and discuss the role that this conceptualisation can play in Information Retrieval systems for reducing online misinformation.

4.1. Keynote Speeches

As part of the Workshop, two keynote speeches were given by **Arkaitz Zubiaga**, on the topic of rumor and claim verification, and **Gabriella Pasi**, on the open issue of considering credibility as a relevance dimension in IR systems.

4.1.1. Towards Automated Fact-checking for Detecting and Verifying Claims

Abstract: Automated fact-checking is a complex task that goes beyond the determination of the veracity of stories. An end-to-end fact-checking pipeline would involve from the initial step of detecting which claims need to be fact-checked (claim checkworthiness detection), then aggregating associated evidence and knowledge, to ultimately summarise all together in a report making a verdict on the veracity of the story. In this talk, I will cover some of my research in these directions when dealing with social media data. I will first briefly discuss research assessing the capacity of untrained users in determining the potential veracity of stories. I will then discuss research in automatically detecting rumours and claims needing verification, as well as in the subsequent steps of collecting crowd stance and evidence towards determining the veracity value of stories. As part of the talk, I will be touching upon the problem of collecting suitable datasets to tackle the task.

- **Arkaitz Zubiaga** is a Lecturer (Assistant Professor) at Queen Mary University of London, UK. He leads the Social Data Science Lab. His research revolves around Social Data Science, interdisciplinary research bridging Computational Social Science and Natural Language Processing. He is particularly interested in linking online data with events in the real world, among others for tackling problematic issues on the Web and social media that can have a damaging effect on individuals or society at large, such as

hate speech, misinformation, inequality, biases and other forms of online harm. *Website:* <http://www.zubiaga.org/>

4.1.2. Credibility and Relevance in Information Retrieval

Abstract: In the field of Information Retrieval, the study of how to ensure access to relevant information with respect to users' information needs has been steadily developing over the last fifty years. It has gone from considering only the topical relevance of documents, to taking into account the concept of popularity in Web search engines, to considering user contextual aspects in personalized search. Nowadays, where we are witnessing the proliferation of misinformation spread through both the Web and social media, a new challenge arises in the IR field: to provide users with information that is also credible. Depending on whether one searches Web pages or social content, depending on the task and the domain for which the search is performed, the concept of credibility understood as an aspect of relevance may change. In this talk, I will therefore present some issues related to credibility assessment in IR, and the problem of constructing environments and datasets for experimental evaluation of approaches that intend to address them.

- **Gabriella Pasi** is Full Professor at the Department of Informatics, Systems, and Communication (DISCo) of the University of Milano-Bicocca. Within DISCo she leads the Information and Knowledge Representation, Retrieval, and Reasoning (IKR3) Lab. Her main research interests are related to Information Retrieval, Recommender Systems, Text Mining, Knowledge Representation and Reasoning, User Modeling. She is also interested in Social Media Analytics and, in particular, in the analysis of User-Generated Content for the study of information dissemination and evolution and information credibility assessment. *Website:* <http://www.ir.disco.unimib.it/people/pasi-gabriella/>

5. Organizing Team

The ROMCIR 2021 organizing team is composed of the following people with respect to their distinct roles.

5.1. Co-chairs

- **Fabio Saracco** is Assistant Professor (RTDa) at IMT School For Advanced Studies since May 2016, where he works in the NETWORKS research unit guided by Prof. Garlaschelli. Fabio's research is devoted to the theoretical development of tools for the analysis of complex networks. Recently these techniques were applied in the context of Online Social Network in the activities of the TOFFEE (TOol for Fighting FakEs) project, funded by the IMT School For Advanced Studies and led by Prof. Rocco De Nicola, and in those of the European Project SoBigData++ (GA. 871042). *Website:* <https://www.imtlucca.it/en/fabio.saracco/>

- **Marco Viviani** is Assistant Professor (RTDb) at the University of Milano-Bicocca, Department of Informatics, Systems, and Communication (DISCo). He works in the Information and Knowledge Representation, Retrieval and Reasoning (IKR3) Lab. He has been co-chair of several special tracks and workshops at international conferences, also related to the assessment of information credibility, and general co-chair of MDAI 2019. His main research activities include Information Retrieval, Social Computing, User Modeling, Trust and Reputation Management. *Website:* <http://www.ir.disco.unimib.it/people/marco-viviani/>

5.2. Publicity Chair

- **Marinella Petrocchi**, Institute of Informatics and Telematics (IIT) – CNR, Pisa, Italy.
Website: <https://www.iit.cnr.it/marinella.petrocchi>

5.3. Program Committee

- **Rino Falcone**, Inst. of Cognitive Sciences and Technologies (ISTC) – CNR, Rome, Italy
- **Carlos A. Iglesias**, Universidad Politécnica de Madrid, Madrid, Spain
- **Petr Knoth**, The Open University, London, UK
- **Udo Kruschwitz**, University of Regensburg, Regensburg, Germany
- **Yelena Mejova**, ISI Foundation, Turin, Italy
- **Preslav Nakov**, Qatar Computing Research Institute, HBKU, Doha, Qatar
- **Symeon Papadopoulos**, Inf. Tech. Inst. (ITI), Thessaloniki, Greece
- **Marinella Petrocchi**, Inst. of Inf. and Telematics (IIT) – CNR, Pisa, Italy
- **Barbara Poblete**, University of Chile, Santiago, Chile
- **Adrian Popescu**, CEA LIST, Gif-sur-Yvette, France
- **Paolo Rosso**, Universitat Politècnica de València, València, Spain
- **Fabio Saracco**, IMT School for Advanced Studies, Lucca, Italy
- **Marco Viviani**, University of Milano-Bicocca, Milan, Italy
- **Xinyi Zhou**, Syracuse University, Syracuse, NY, USA
- **Arkaitz Zubiaga**, Queen Mary University of London, London, UK

Acknowledgments

We would like to thank the authors of the submitted articles for their interest in the considered problem, the Keynote Speakers for the interest aroused in new research directions, and the members of the Program Committee for their valuable contribution to the success of the ROMCIR 2021 Workshop.

F.S. acknowledges also support from the European Project SoBigData++ (GA. 871042) and the PAI (*Progetto di Attività Integrata*) project TOFFeE, funded by the IMT School For Advanced Studies Lucca.

References

- [1] G. Pasi, M. Viviani, Information credibility in the social web: Contexts, approaches, and open issues, arXiv preprint arXiv:2001.09473 (2020).
- [2] M. Viviani, G. Pasi, Credibility in social media: opinions, news, and health information—a survey, WIREs: Data Mining and Knowl. Discovery 7 (2017) e1209.
- [3] D. M. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, et al., The science of fake news, Science 359 (2018) 1094–1096.
- [4] Y. Wu, E. W. Ngai, P. Wu, C. Wu, Fake online reviews: Literature review, synthesis, and directions for future research, Decision Support Systems (2020) 113280.
- [5] F. Tagliabue, L. Galassi, P. Mariani, The “pandemic” of disinformation in covid-19, SN Comprehensive Clinical Medicine 2 (2020) 1287–1289.
- [6] C. Wardle, H. Derakhshan, Information disorder: Toward an interdisciplinary framework for research and policy making, Council of Europe Report 27 (2017).
- [7] B. Carminati, E. Ferrari, M. Viviani, Security and trust in online social networks, Synthesis Lectures on Information Security, Privacy, & Trust 4 (2013) 1–120.
- [8] G. Eysenbach, Credibility of health information and digital media: New perspectives and implications for youth, in: M. M. Metzger, A. J. Flanagin (Eds.), Digital Media, Youth, and Credibility, The MIT Press, 2008, pp. 123–154.
- [9] M. J. Metzger, A. J. Flanagin, Credibility and trust of information in online environments: The use of cognitive heuristics, Journal of pragmatics 59 (2013) 210–220.
- [10] J. Thorne, A. Vlachos, O. Cocarascu, C. Christodoulopoulos, A. Mittal, The Fact Extraction and VERification (FEVER) shared task, arXiv preprint arXiv:1811.10971 (2018).