Watch ’n’ Check: Towards a Social Media Monitoring Tool to Assist Fact-Checking Experts

Assunta Cerone*, Elham Naghizade†, Falk Scholer†, Devi Mallal‡, Russell Skelton‡, Damiano Spina†
*Politecnico di Torino
Turin, Italy
Email: assunta.cerone@studenti.polito.it
†RMIT University
Melbourne, Australia
Email: {e.naghizade, falk.scholer, damiano.spina}@rmit.edu.au
‡RMIT ABC Fact Check
Melbourne, Australia
Email: mallal.devi@abc.net.au, russell.skelton@rmit.edu.au

***Abstract*—We present an ongoing collaboration between computer science researchers and fact-checking experts in a broadcast corporation to develop *Watch ’n’ Check*, a social media monitoring tool that assists fact-checkers to detect and target misinformation online. The lean methodology followed in our collaboration has helped us to better understand how information access tools can assist fact-checking experts. We report initial results and discuss our plan for further development, as well as the open challenges identified so far.**

***Index Terms*—social media monitoring, fact-checking, information access tool**

I. Introduction

Societies used to rely on media experts and journalists to take a wide variety of decisions with regard to what, when and how to broadcast a piece of information as news. With the advent of social media platforms, where content can be produced and distributed by anyone, and be promoted automatically, there has been a tremendous shift in how and what we consume as news. Recent studies show that the number of U.S. adults that—at least occasionally—get news on online platforms has increased from 49% in 2012 to 68% in 2018.

A major drawback of modern and free media outlets is their widespread (mis)use to disseminate misinformation[1] with devastating economic, social, and political outcomes. In January 2020, for instance, bots and trolls have been spreading disinformation exaggerating the role of arson to undermine the link between bushfires in Australia and climate change [1]. The World Health Organization (WHO) has referred to the problem of large amount of misinformation spread during the COVID-19 pandemic as an “infodemic”[2] [2]. Verifying information published in social media is therefore of great importance, in order to avoid further costs on society.

With the surge in the spread of misinformation in social media, there has been several studies that aim to detect or mitigate such spread through analyzing and characterizing the news content, source, users, network structure, or a combination of these factors [3]–[6]. In this scenario, fact-checking organizations play a key role in the monitoring of news and social media, identifying, and verifying disruptive claims. However, it is yet unclear how information access tools—and, in particular, social monitoring tools—would be integrated into the workflows and processes carried out by fact checkers. In this paper, we present a prototype of *Watch ’n’ Check*, a tool that can be used by fact-checking experts to facilitate the access and monitoring of social media platforms such as Twitter.

Our *Watch ’n’ Check* prototype has been developed in collaboration with fact-checking experts from the Australian Broadcasting Corporation (ABC).[3] The aim of our collaboration is to identify key functionalities which would inform the design and development of an information access tool to assist fact checkers with verifying information in social media.[4]

The tool aims to complement the fact checkers daily work, by assisting them in the identification and targeting of misleading claims. The prototype developed so far sheds some light on the methodology to be used to develop *Watch ’n’ Check*. We believe there is potential to build a tool to: (i) access large information streams published in real-time; (ii) enable an efficient identification of trending topics, and facilitate monitoring of changes in news propagation dynamics; and (iii) to improve the impartiality in the process of targeting the statements/news to be checked.

The remainder of the paper is organized as follows. Section II summarizes the related work. Section III describes

[1] We acknowledge the distinctions between misinformation, rumor, disinformation, and fake news. In this paper, however, we use the term *misinformation* to refer to them all, since the fact-checking process aims to determine the veracity and correctness of information, and the degrees of veracity does not affect the described methodology.

[2] https://www.who.int/dg/speeches/detail/director-general-s-remarks-at-the-media-briefing-on-2019-novel-coronavirus---8-february-2020

[3] https://www.abc.net.au/news/factcheck/

[4] Currently, *Watch ’n’ Check* is only accessible by the authors of the paper. The analyses carried out so far are based on aggregated data and do not expose sensitive personal information.

the methodology—based on quick iterations—being used to develop the tool. Section IV details the results of each of the iterations. We discuss our findings and describe the current limitations of the prototype in Section V. Finally, Section VI concludes the work and identifies challenges that remain open.

II. Related Work

An abundance of sophisticated automated tools and machine learning models have been proposed to accurately detect and consequently mitigate the spread of fake news in social media [3], e.g., by developing fact-checking URL recommender systems and text generation models [5], [7], [8]. There have been also attempts to address the fact-checking problem from an automatic point of view [9] include automatic systems for check-worthiness prediction [10], [11] or truthfulness detection/credibility assessment [4], [6], [12]–[16].

Recent work has studied the phenomenon of spread of fake news is social media. Vosoughi, Roy, and Aral [17] have found that lies spread faster than the truth, and have observed evidence of the relation between emotion and veracity. Shu et al. [18] released FakesNewsNet, a fake news data repository containing news content, social context, and dynamic information that can be used to analyze the spread of fake news in social media.

The problem of verifying statements has also studied from a crowdsourcing perspective, by analyzing the impact of truthfulness scales and workers' background when collecting truthfulness judgments [14], [19]–[24].

However, little attention has been paid to better understand how fact-checking experts can integrate information access tools to inform the processes that are currently being used to verify information in social media.

Most of the social media monitoring tools proposed in the literature, such as RAPID [25], VATAS [26], or ORMA [27] aim to provide insights for other scenarios than fact-checking, such as marketing, customer experience, or online reputation monitoring. Significant effort is being made to analyze COVID-19 information on social media, and linking to data from external fact-checking organizations to quantify the spread of misinformation [2], [28]–[31].

III. Methodology

We firstly describe the lean methodology adopted in our collaboration with fact checkers to develop *Watch 'n' Check*. Then, we briefly describe the Twitter dataset and the topics identified together with the experts during our collaboration.

Lean Methodology: Figure 1 illustrates the *lean* methodology [32] that was followed to develop the *Watch 'n' Check* tool. Each iteration consists of four phases. In the first phase, the researchers—all of them with a Computer Science background—meet with the fact-checking experts—with a journalism and communication background—to identify key functionality, i.e., define the essential features that the tool must have. In the second phase, researchers have brainstorming sessions to understand where the "low hanging fruit" is, and to identify the most cost-effective way to apply text

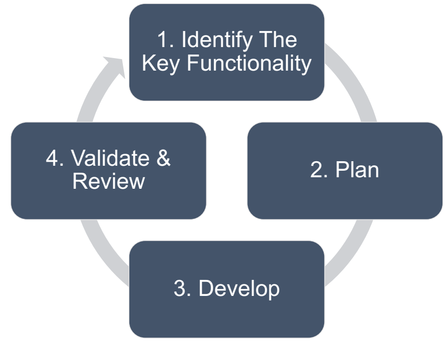

Fig. 1. Iterative methodology used to collaborate with fact-checking experts.

TABLE I
TOPICS AND ASSOCIATED KEYWORDS USED FOR OUR COLLABORATION WITH FACT-CHECKING EXPERTS.

Topic	Keywords
Bushfires in Australia	`bushfires`, `climate change`, `arson`
COVID-19	`coronavirus`, `covid`, `vaccine`

analytics and data visualization tools to develop a Minimal Viable Product (MVP) to show to the experts, which is the third phase. Finally, the MVP is validated by the experts, who provide feedback about the implemented features, which would also inform the starting point of the next iteration.

Dataset: We started collecting tweets from December 1, 2019. Tweets are collected by consuming Twitter's Sample Stream V1 API,[5] using the `Twitter4j` client.[6] The API supports the collection of around 1% of publicly available tweets, as they happen. Tweets that are non-empty and in English language—as labeled by the API—are then indexed using Elasticsearch.[7] As at May 1, 2020, the index contains a total of 182.1M tweets, with an average of 1.2M tweets per day.

Topics: There were two main topics that the fact-checking experts were particularly interested in exploring with *Watch 'n' Check*: the extreme bushfires that occurred during December 2019 and January 2020 in Australia, and the COVID-19 pandemic. Both major events triggered the spread of a significant amount of misinformation (and disinformation) on social media [33]. Table I lists the keywords defined together with the experts to use as case studies for our iterations. For instance, we included the keyword `arson` as the fact checkers in our team identified that it was being used frequently in counter-narratives to climate change being a main driver of the bushfires in Australia.

IV. Preliminary Results

So far, we have performed three iterations, which are described below. Each iteration spanned about three weeks.

[5] https://developer.twitter.com/en/docs/labs/sampled-stream/overview
[6] http://twitter4j.org/en/
[7] https://www.elastic.co/

A. First Iteration – Identification of the Target Platform

The first question raised by the experts was about the access to public data published in social media platforms. Misinformation can be spread through multiple channels, including Facebook groups, Instagram conversations, or Twitter posts, among others. We identified Twitter as our initial source to explore, as (i) it is one of the most important channels to spread information online, and (ii) it provides an API to extract a representative sample of the published information as it is released.

Proposed Functionality: We planned and developed a console-based Python script that provided a simple but comprehensive way of inspecting an indexed collection of tweets obtained by the Twitter API.

The first proposed feature aimed to perform a quantitative real-time analysis on specific topics. Given a keyword, it computed and allowed the checking of:

- Number of tweets per month;
- Number of tweets per location;
- Number of tweets per user; and
- Number of tweets per day.

Validation & Review: At this stage, fact checkers were invited to give their feedback on the current system in order to confirm its usefulness in their fact-checking process, and suggest new features. In this first phase, the following requirements were identified:

- Extend the analysis to a phrase or a set of words, instead of a single word. Fact checkers were not only interested in the analysis of keywords (e.g., hashtags), but also in detecting and tracking entire statements or phrases (e.g., politicians' claims).
- Develop a user-friendly and compact visualization of the data. The experts would gain more insights with a graphical representations of the aggregated data (e.g., understand how the tweets about bushfires evolve over time in order to correlate this with external events).
- Provide access to tweet instances to have an understanding of their content. Besides descriptive statistics that summarize an aggregated sample of tweets, fact checkers were interested in inspecting a sample of the textual content in the tweets, which could then potentially be used to perform more in-depth manual analysis.

B. Second Iteration – Visualizing Trends

The requirements identified in the previous iteration were used as a starting point for both the improvement of the current functionality, and the identification of new ones. In particular, we evaluated again the possibility to collect data from other platforms such as Facebook and Instagram. However, given the limitations of obtaining public data on those platforms, we decided to focus on developing the *Watch 'n' Check* over Twitter data.

Therefore, we planned the following features to be developed during the second iteration:

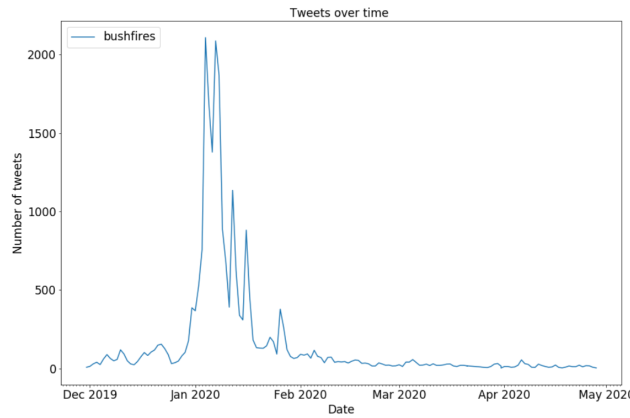

Fig. 2. Frequency over time of tweets in the collection that contain the keyword `bushfires`.

- Visualization of the number of tweets containing a keyword over time.
- Computing the most frequent unigrams, bigrams, and trigrams as a mechanism to provide an overview over tweet texts.

Proposed Functionality: The first new functionality proposed is the visualization of the frequency of tweets related to a specific keyword/phrase over time. Figure 2, for instance, illustrates the frequency of tweets containing the keyword `bushfires` in our collection from December 2019 to April 2020.

The second functionality aims to provide a simple mechanism to explore the most co-occurrent unigrams, bigrams, and trigrams for a given keyword in the tweet texts. It supports a dynamic choice of the period of time to analyze, and computes n-grams in that period. Figure 3 shows the graphs by *Watch 'n' Check* with the most co-occurrent n-grams for the keyword `bushfires` in the four indexed months of 2020.

Validation & Review: The first feature of this second iteration allows the experts to verify how a certain keyword evolves over time on Twitter. They also found it interesting to understand how a trend for a given keyword compares with the trend for other keywords/phrases. This can guide experts *a priori* in focusing their attention on the analysis of some topics instead of others, or it can be used as a tool to retrospectively validate the choice to focus on fact-checking certain content.

The n-gram graphs provide an approximate overview of what the topical content of the tweets. While it is useful to have a general idea, the fact-checking experts need to have a more detailed introspection on the content of the tweets, to check their correlation with some news (or fake news), and to potentially track their spread. Another aim of the fact checkers is to gain an understanding about the popularity of some claims or user profiles rather than keywords.

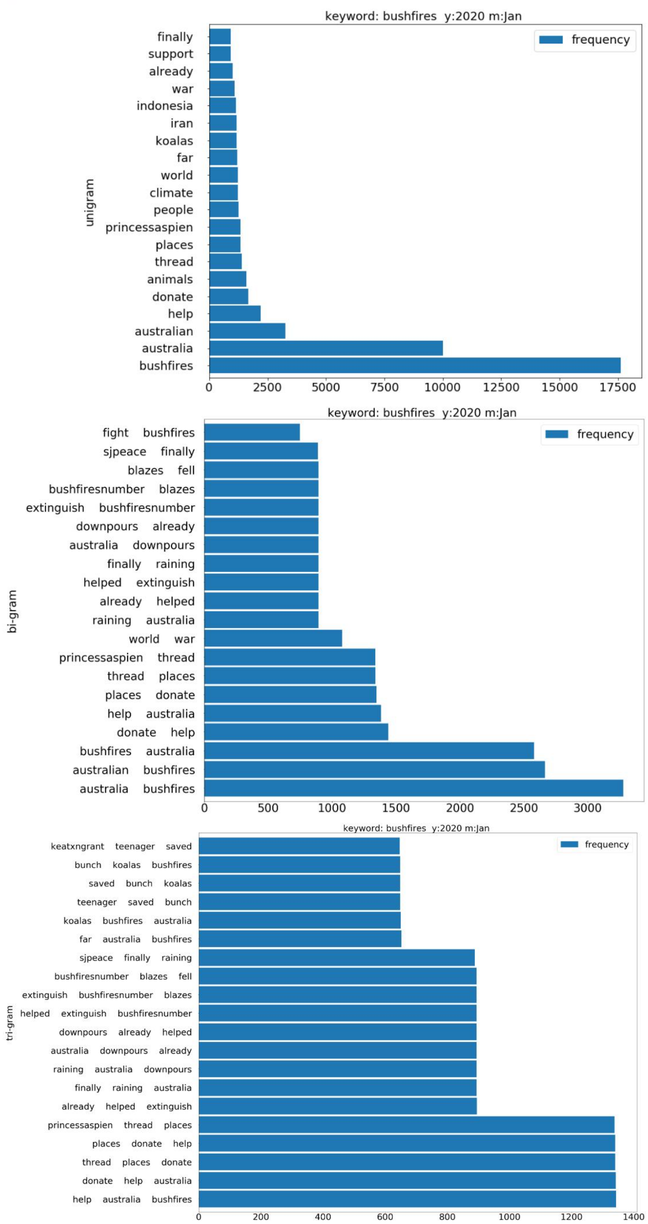

Fig. 3. Frequency of n-grams which co-occur with the keyword `bushfires`.

C. Third Iteration – Towards a User-friendly Interface

After two iterations using a console-based interface operated by the researchers, Jupyter Notebook[8] has been identified as a tool that would allow to quickly provide a more friendly user interface. Another key functionality that was identified is the relative comparison of multiple keywords over time, as it would provide an easy and immediate way to analyze the lifespan of different keywords in a single visualization.

Proposed Functionality: In order to provide a user interface that would allow experts to generate their own analyses, we opted to use Jupyter Notebook. Although not ideal—a web front-end is planned for future work as described in Section V—it provided a flexible and quick way to show results in a more friendly manner. The new interface allows access to all the functionality from a browser with access to the intranet.

The second new functionality of the the tool provides a comparison of the trends of multiple keywords simultaneously. Figure 4 illustrates this functionality for the keywords `bushfires`, `climate change`, `coronavirus`, and `vaccine`. As different keywords may generate curves with substantial changes in the range of frequencies, a graph with a logarithmic scale is also generated.

Finally, we also developed the functionality to dynamically extract a random sample of tweets from a chosen period of time. The sample is then exported to a CSV file, which includes the following columns: tweet id, user id, content of the tweet, and timestamp. Table II shows some (de-identified) tweets extracted for the keywords `arson` and `coronavirus`.

Validation & Review: The comparison of the frequencies over time provides valuable insights for the experts. In fact, one of the graphs generated by the tool has been included in their weekly fact-checking newsletter. An interesting obser-

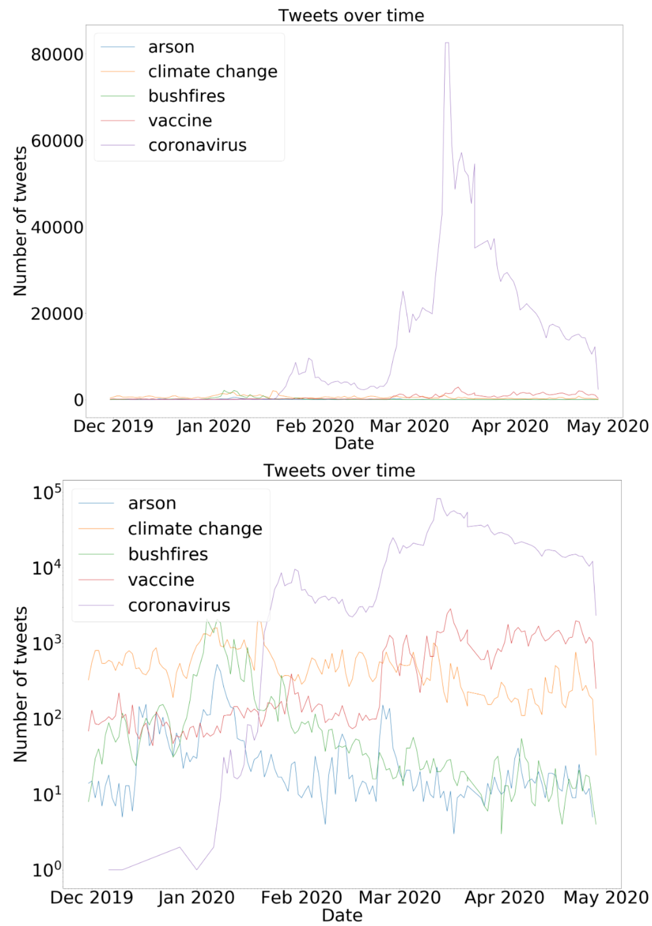

Fig. 4. Comparison of the frequency over time of tweets in the collection that contains the specified keywords.

[8]https://jupyter.org/

TABLE II

EXAMPLE OF TWEETS EXTRACTED FROM A SAMPLE FOR THE KEYWORDS `arson` AND `coronavirus`.

Keyword	Tweet Content
`arson`	@USER The fire in Australia is arson. The Chinese virus is eating bats or lab made. Its the people not God. RT @USER: . Climate scam – 24 charged in Australian bush fires arson arson is the cause of the Australian bush fires RT @USER: The fires were really arson. It was miraculous they were rained out. It was #prayer they can't explain it.
`coronavirus`	RT @USER: Just in: @realdonaldtrump donates his quarterly salary to @hhsgov to help fight coronavirus. Thank you, President Trump! RT @USER: In the United States it's easier to get an abortion than get tested for coronavirus. RT @USER: As a result of president @realdonaldtrump's leadership, every state lab in the country can now conduct coronavirus testing.

vation at this stage was that the linear scale version of the graph was preferred over the logarithmic scale counterpart. This suggests that, although some visualization techniques may provide a clearer representation of skewed data, these may be harder to interpret.

The functionalities of *Watch 'n' Check* integrated are into a Jupyter Notebook seem to appear clear to experts. However, we acknowledge that is not optimal, as the user could inadvertently change the structure of the code and compromise the analysis itself. Furthermore, the fact-checking process analyzes the truthfulness of news and claims on a daily/weekly basis. This highlights the need to have a real-time application and create an easily accessible user interface for fact checkers.

V. DISCUSSION

Findings: We described the ongoing work to develop *Watch 'n' Check*, an information access tool that aims to assist fact checkers during the process of verifying information in social media. Using a lean methodology, and with the collaboration of fact-checking experts, we were able to identify, plan, develop, and validate the key functionalities of our tool. We found that providing mechanisms to track keywords over time, and co-occurrent n-grams, are as important as providing access to instances of data (i.e., tweets containing the keywords of interest). We also found that there is a compromise between informativeness and comprehensibility when visualizing the trends. For instance, graphs in linear scale may be preferred against their logarithmic scale counterpart, as they are easier to understand.

Review sessions in each iteration offer insights in terms of identifying research gaps. For instance, the sharp drop in the number of `coronavirus`-related tweets in mid-March (Figure 4) was first identified as a potential issue in the data crawling process, however, the fact-checking experts mentioned other plausible explanations for this phenomenon such as weariness of the audience, or seeking information from alternative sources, which has also been observed in other areas. Exploring such hypotheses requires a more in-depth analysis which is beyond the scope of the current research, yet is insightful for future research endeavors.

Limitations: The current version of *Watch 'n' Check* includes a number of limitations. The current prototype relies on the sample of tweets collected via the Twitter Stream API—which exposes about 1 percent of publicly available tweets—and is further filtered by language. Therefore, *Watch 'n' Check* can be used as a complementary tool to help identify relevant information, but experts will still need to access the original platform to refine their analyses. Moreover, having access to multiple social media platforms has been identified as a key functionality. However, other social media platforms such as Instagram and Facebook have more restrictive access to public data. Finally, the current user interface provided through Jupyter Notebook was used for the purpose of offering a Minimal Viable Product, as an easy way to let the experts validate the functionalities provided. The development of a web front-end is part of our immediate future work, as well as the inclusion of an analysis of hashtags and URLs contained in the tweets.

VI. CONCLUSIONS AND FUTURE WORK

In this paper we present a prototype of *Watch 'n' Check*, a social media monitoring tool that has been designed and developed to specifically help fact checkers with their task of verifying the veracity of news and claims in social media platforms. We achieved this through following a lean methodology and incorporating the fact checkers' feedback into each iteration of our development process. The source code of the *Watch 'n' Check* prototype is publicly available.[9]

Watch 'n' Check represents the base for more advanced and accurate analysis that can improve the connection between fact-checking and information propagation in social media. Current analyses can be extended beyond the information content by analyzing the social network structure and the users who engage with this content. In this way, fact checkers can effectively analyze trends and communities and their associations, which would inform their process of targeting information in social media that needs to be verified.

Watch 'n' Check filters tweets by matching the specified keyword or phrase. However, a semantic representation of the *topic* —e.g., by the use of word embeddings—would enhance

[9] https://github.com/rmit-ir/watch-n-check

the analysis. To this aim, we plan to incorporate automatic keyword extraction [34] and topic detection [35], [36] methods, so if the experts filter tweets about `coronavirus`, the tool would also filter tweets that are semantically related (e.g., tweets that include `covid`). In addition, there is evidence of the relationship between emotions and the spread of misinformation [17], and recent work has shown the effectiveness of modeling emotions for automatic identification of false information in social media content [37], [38]. Future work includes enhancing the analysis with automatic classification of emotions or fake news spreaders [39]–[41].

There have been significant advances in research to address the problems of automatic bot and fake news detection. A challenge that remains open is to understand how these sophisticated tools can be integrated in the fact-checking process in an effective and transparent way, to support the experts in their daily work.

Finally, one other challenge in designing information access tools for fact checkers—and for any given group of experts in general—consists of having a clear way to explain the output of the system, e.g., how a conclusion is made and with how much confidence, to ensure the main objective of such system is reached, i.e., empowering experts through providing timely and accurate information. Besides continuing using the lean methodology described in this paper, we plan to perform user studies to better understand how *Watch 'n' Check* can effectively incorporated into the fact-checking process performed by experts in their daily work.

Acknowledgment

This work is partially supported by the Australian Research Council (project nr. DE200100064).

References

[1] T. Graham and T. R. Keller, "Bushfires, Bots and Arson Claims: Australia Flung in the Global Disinformation Spotlight," *The Conversation*, 2020, Accessed: 29-05-2020. [Online]. Available: https://theconversation.com/bushfires-bots-and-arson-claims-australia-flung-in-the-global-disinformation-spotlight-129556

[2] F. Alam, S. Shaar, A. Nikolov, H. Mubarak, G. D. S. Martino, A. Abdelali, F. Dalvi, N. Durrani, H. Sajjad, K. Darwish *et al.*, "Fighting the COVID-19 Infodemic: Modeling the Perspective of Journalists, Fact-Checkers, Social Media Platforms, Policy Makers, and the Society," *arXiv preprint arXiv:2005.00033*, 2020.

[3] K. Sharma, F. Qian, H. Jiang, N. Ruchansky, M. Zhang, and Y. Liu, "Combating Fake News: A Survey on Identification and Mitigation Techniques," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 10, no. 3, pp. 1–42, 2019.

[4] P. Atanasova, P. Nakov, L. Màrquez, A. Barrón-Cedeño, G. Karadzhov, T. Mihaylova, M. Mohtarami, and J. Glass, "Automatic Fact-Checking Using Context and Discourse Information," *Journal of Data and Information Quality*, vol. 11, no. 3, 2019.

[5] N. Vo and K. Lee, "The Rise of Guardians: Fact-checking URL Recommendation to Combat Fake News," in *Proceedings of SIGIR'18*, 2018, pp. 275–284.

[6] K. K. Popat, "Credibility Analysis of Textual Claims with Explainable Evidence," Ph.D. dissertation, Saarland University, Saarbrücken, 2019.

[7] N. Vo and K. Lee, "Learning from Fact-checkers: Analysis and Generation of Fact-checking Language," in *Proceedings of SIGIR'19*, 2019, pp. 335–344.

[8] D. You, N. Vo, K. Lee, and Q. LIU, "Attributed Multi-Relational Attention Network for Fact-Checking URL Recommendation," in *Proceedings of CIKM'19*, 2019, p. 1471–1480.

[9] A. Vlachos and S. Riedel, "Fact Checking: Task Definition and Dataset Construction," in *Proceedings of the ACL 2014 Workshop on Language Technologies and Computational Social Science*, 2014, pp. 18–22.

[10] P. Gencheva, P. Nakov, L. Màrquez, A. Barrón-Cedeño, and I. Koychev, "A Context-Aware Approach for Detecting Worth-Checking Claims in Political Debates," in *Proceedings of the International Conference Recent Advances in Natural Language Processing, (RANLP'17)*, 2017, pp. 267–276.

[11] S. Vasileva, P. Atanasova, L. Màrquez, A. Barrón-Cedeño, and P. Nakov, "It Takes Nine to Smell a Rat: Neural Multi-Task Learning for Check-Worthiness Prediction," in *Proceedings of Recent Advances in Natural Language Processing (RANLP'19)*, 2019, pp. 1229–1239.

[12] T. Mihaylova, G. Karadjov, P. Atanasova, R. Baly, M. Mohtarami, and P. Nakov, "SemEval-2019 Task 8: Fact Checking in Community Question Answering Forums," in *Proceedings of SemEval*, 2019, pp. 860–869.

[13] P. Nakov, A. Barrón-Cedeno, T. Elsayed, R. Suwaileh, L. Màrquez, W. Zaghouani, P. Atanasova, S. Kyuchukov, and G. Da San Martino, "Overview of the CLEF-2018 CheckThat! Lab on Automatic Identification and Verification of Political Claims," in *Proceedings of CLEF'18*, 2018, pp. 372–387.

[14] T. Elsayed, P. Nakov, A. Barrón-Cedeño, M. Hasanain, R. Suwaileh, G. Da San Martino, and P. Atanasova, "Overview of the CLEF-2019 CheckThat! Lab: Automatic Identification and Verification of Claims," in *Proceedings of CLEF'19*, 2019, pp. 301–321.

[15] J. Kim, D. Kim, and A. Oh, "Homogeneity-Based Transmissive Process to Model True and False News in Social Networks," in *Proceedings of WSDM'19*, 2019, p. 348–356.

[16] A. Barrón-Cedeño, T. Elsayed, P. Nakov, G. Da San Martino, M. Hasanain, R. Suwaileh, and F. Haouari, "CheckThat! at CLEF 2020: Enabling the Automatic Identification and Verification of Claims in Social Media," in *Proceedings of ECIR'20*, 2020, pp. 499–507.

[17] S. Vosoughi, D. Roy, and S. Aral, "The Spread of True and False News Online," *Science*, vol. 359, no. 6380, pp. 1146–1151, 2018.

[18] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, "FakeNewsNet: A Data Repository with News Content, Social Context, and Spatiotemporal Information for Studying Fake News on Social Media," *Big Data*, vol. 8, no. 3, pp. 171–188, 2020.

[19] T. Kriplean, C. Bonnar, A. Borning, B. Kinney, and B. Gill, "Integrating On-demand Fact-checking with Public Dialogue," in *Proceedings of CSCW'14*, 2014, pp. 1188–1199.

[20] A. Zubiaga and H. Ji, "Tweet, but Verify: Epistemic Study of Information Verification on Twitter," *Social Network Analysis and Mining*, vol. 4, no. 1, pp. 1–12, 2014.

[21] K. Roitero, M. Soprano, B. Portelli, D. Spina, V. Della Mea, G. Serra, S. Mizzaro, and G. Demartini, "The COVID-19 Infodemic: Can the Crowd Judge Recent Misinformation Objectively?" in *Proceedings of CIKM'20*, 2020.

[22] K. Roitero, M. Soprano, S. Fan, D. Spina, S. Mizzaro, and G. Demartini, "Can The Crowd Identify Misinformation Objectively? The Effects of Judgment Scale and Assessor's Background," in *Proceedings of SIGIR'20*, 2020, pp. 439–448.

[23] D. La Barbera, K. Roitero, G. Demartini, S. Mizzaro, and D. Spina, "Crowdsourcing Truthfulness: The Impact of Judgment Scale and Assessor Bias," in *Proceedings of ECIR'20*, 2020, pp. 207–214.

[24] K. Roitero, G. Demartini, S. Mizzaro, and D. Spina, "How Many Truth Levels? Six? One Hundred? Even more? Validating Truthfulness of Statements via Crowdsourcing," in *CIKM Workshop on Rumours and Deception in Social Media (RDSM'18)*, 2018.

[25] K. H. Lim, S. Jayasekara, S. Karunasekera, A. Harwood, L. Falzon, J. Dunn, and G. Burgess, "RAPID: Real-time Analytics Platform for Interactive Data Mining," in *Proceedings of ECML'18*, 2018, pp. 649–653.

[26] D. U. Patton, P. Blandfort, W. R. Frey, R. Schifanella, K. McGregor, and S.-F. U. Chang, "VATAS: An Open-Source Web Platform for Visual and Textual Analysis of Social Media," *Journal of the Society for Social Work and Research*, vol. 11, no. 1, pp. 133–155, 2020.

[27] J. Carrillo-de Albornoz, E. Amigó, D. Spina, and J. Gonzalo, "ORMA: A Semi-automatic Tool for Online Reputation Monitoring in Twitter," in *Proceedings of ECIR'14*, 2014, pp. 742–745.

[28] R. Gallotti, F. Valle, N. Castaldo, P. Sacco, and M. D. Domenico, "Assessing the Risks of "Infodemics" in Response to COVID-19 Epidemics," *arXiv preprint arXiv:2004.03997*, 2020.

[29] M. Cinelli, W. Quattrociocchi, A. Galeazzi, C. M. Valensise, E. Brugnoli, A. L. Schmidt, P. Zola, F. Zollo, and A. Scala, "The COVID-19 Social Media Infodemic," *arXiv preprint arXiv:2003.05004*, 2020.
[30] K.-C. Yang, C. Torres-Lugo, and F. Menczer, "Prevalence of Low-Credibility Information on Twitter During the COVID-19 Outbreak," *arXiv preprint arXiv:2004.14484*, 2020.
[31] J. Bullock, A. Luccioni, K. H. Pham, C. S. N. Lam, and M. Luengo-Oroz, "Mapping the Landscape of Artificial Intelligence Applications against COVID-19," *arXiv preprint arXiv:2003.11336*, 2020.
[32] E. Reis, *The Lean Startup*. Portfolio Penguin, 2011.
[33] C. Zappone, "#ArsonEmergency: How 'Fake News' Created an Information Crisis about the Bushfires," *The Sydney Morning Herald*, 2020, Accessed: 29-05-2020. [Online]. Available: https://www.smh.com.au/world/oceania/arsonemergency-how-fake-news-created-an-information-crisis-about-the-bushfires-20200211-p53zma.html
[34] D. Spina, J. Gonzalo, and E. Amigó, "Discovering Filter Keywords for Company Name Disambiguation in Twitter," *Expert Systems with Applications*, vol. 40, no. 12, pp. 4986–5003, 2013.
[35] D. Spina, J. Gonzalo, and E. Amigó, "Learning Similarity Functions for Topic Detection in Online Reputation Monitoring," in *Proceedings of SIGIR'14*, 2014, pp. 527–536.
[36] D. M. Blei, "Probabilistic Topic Models," *Communications of the ACM*, vol. 55, no. 4, pp. 77–84, 2012.
[37] B. Ghanem, P. Rosso, and F. Rangel, "An Emotional Analysis of False Information in Social Media and News Articles," *ACM Transactions on Internet Technology (TOIT)*, vol. 20, no. 2, pp. 1–18, 2020.
[38] A. Giachanou, P. Rosso, and F. Crestani, "Leveraging Emotional Signals for Credibility Detection," in *Proceedings of SIGIR'19*, 2019, pp. 877–880.
[39] F. Rangel, A. Giachanou, B. Ghanem, and P. Rosso, "Overview of the 8th Author Profiling Task at PAN 2020: Profiling Fake News Spreaders on Twitter," in *CLEF 2020 Labs and Workshops, Notebook Papers*. CEUR Workshop Proceedings, 2020. [Online]. Available: CEUR-WS.org
[40] X. Duan, E. Naghizade, D. Spina, and X. Zhang, "RMIT at PAN-CLEF 2020: Profiling Fake News Spreaders on Twitter," in *CLEF 2020 Labs and Workshops, Notebook Papers*. CEUR Workshop Proceedings, 2020. [Online]. Available: CEUR-WS.org
[41] K. Shu, S. Wang, and H. Liu, "Understanding user profiles on social media for fake news detection," in *2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, 2018, pp. 430–435.