

Four years of fake news: A quantitative analysis of the scientific literature

by Nicola Righetti

Abstract

Since 2016, “fake news” has been the main buzzword for online misinformation and disinformation. This term has been widely used and discussed by scholars, leading to hundreds of publications in a few years. This report provides a quantitative analysis of the scientific literature on this topic by using frequency analysis of metadata and automated lexical analysis of 2,368 scientific documents retrieved from Scopus, a large scientific database, mentioning “fake news” in the title or abstract.

Findings show that until 2016 the number of documents mentioning the term was less than 10 per year, suddenly rising from 2017 and steadily increasing in the following years. Among the most prolific countries are the U.S. and European countries such as the U.K., but also many non-Western countries such as India and China. Computer science and social sciences are the disciplinary fields with the largest number of documents published. Three main thematic areas emerged: computational methodologies for fake news detection, the social and individual dimension of fake news, and fake news in the public and political sphere. There are 10 documents with more than 200 citations, and two papers with a record number of citations.

Contents

[Introduction](#)

[Methods and research questions](#)

[Analysis](#)

[Discussion and conclusions](#)

Introduction

Since the 2016 American presidential election, “fake news” has become the catch-all buzzword for everything related to online misinformation and disinformation, both in journalistic and political discourse (Egelhofer and Lechele, 2019; Egelhofer, *et al.*, 2020; Farkas and Schou, 2018), and scholarly research. Defined as “fabricated information that mimics news media content in form but not in organizational process or intent” [1] or “news articles that are intentionally and verifiably false, and could mislead readers” [2], this word has raised noticeable debate among scholars. Some have criticized the term as it “doesn’t begin to describe the complexity of the different types of misinformation (the inadvertent sharing of false information) and disinformation (the deliberate creation and sharing of information known to be false)” (Wardle, 2017), while others “have retained it because of its value as a scientific construct, and because its political salience draws attention to an important subject” [3].

During the past few years, scholars have dug into online misinformation and disinformation processes, shedding light on the complexity of the phenomenon (*e.g.*, Giglietto, *et al.*, 2019; Marwick and Lewis, 2017; Venturini, 2019), and developing and advocating a more nuanced conceptual and lexical apparatus (*e.g.*, Jack, 2017; Tandoc, *et al.*, 2018;

TaSC, 2020). Nonetheless, “fake news” has never ceased to be a popular term to refer — in a manner as effective as it is simple and, maybe, sometimes simplistic — to the landscape of online problematic information.

Four years on from the consecration of “fake news” in the chronicles of newspapers and academic debates, this report aims at providing readers with a quantitative overview of the scientific literature on the topic.

Methods and research questions

Scientific literature mentioning the term “fake news” was collected from the scientific database Scopus, “a source-neutral abstract and citation database curated by independent subject matter experts”. Scopus includes over 25,100 titles from more than 5,000 international publishers for over 77.8 million records (Scopus, 2020). The wide range of sources included in the database and its multidisciplinary character, make it appropriate for a literature review on a deeply multidisciplinary topic like fake news.

All the document records mentioning the keyword “fake news” in the title or abstract, written in English and published up to 2020, were searched and downloaded along with bibliographic information such as abstract, authors, title, publication name, year of publication, country, and keywords.

The general analytic approach included frequency analysis of bibliographical information and text mining analysis of titles and abstracts. More specifically, the research questions, and methods used to answer them, were as follows:

RQ1: How many documents on “fake news” have been published over time? Is there a growing trend in publications on the topic?

This question was answered by calculating the total number of documents per year and analyzing the related time series. Since the absolute number of publications does not account for a possible general growth in scientific productivity, the number of publications has been further weighted for another clearly popular topic in communication science (Dalmajer, *et al.*, 2021), namely “social media”. Additional descriptive statistics related to sources, authors, citations, and references have been reported. To contextualize the growth of the phenomenon, statistics on worldwide online search trends (from Google Trends: <https://trends.google.com>) and online news media articles mentioning the topic (news media sources in English from Media Cloud: <https://mediacloud.org>) have been provided.

RQ2: What is the geographical distribution of the documents?

This question was answered by calculating the number of documents by authors’ affiliation country, information provided by Scopus metadata.

RQ3: Which are the disciplinary areas of the documents?

This question was answered through frequency analysis of the “subject area” field provided by Scopus metadata.

RQ4: Which are the main topics of the documents, and how have they changed over time?

Topics were analyzed through frequency analysis of keywords associated with each record, and lexical cluster analysis of titles and abstracts performed by using the Reinert descending hierarchical classification algorithm (Ratinaud, 2008; Reinert, 1990). The cluster analysis was performed on 2,172 out of 2,368 (92 percent) documents including an abstract. To identify more fine-grained topics and their change over time, a structural topic modeling approach was implemented (Roberts, *et al.*, 2019).

RQ5: Which are the most cited documents?

This question was answered by analyzing the number of citations reported by Scopus (citation metrics might vary between databases), and by providing an overview of the most popular papers on the topic up to now.

Analysis

A total of 9,015 documents mentioned the term “fake news” somewhere in their full text, and 2,368 documents mentioned it in the title or abstract and were therefore analyzed to answer the research questions.

Quantity of publications and trends

The 2,368 documents were written by 5,060 authors and published on 1,225 different sources, include a total of 81,257 references, have 7.3 average citations per document and about two citations per year per document. Out of 2,368 documents, 1,055 are journal articles and 828 are conference papers, together representing 1,883 records or 80 percent of the set of documents mentioning “fake news” in the title (1,220, 51.5 percent), or abstract (1,148, 48.5 percent). Other documents included are reviews (127 documents), conference reviews (75), book chapters (73), editorials (72), notes (59), books (43), letters (17), short surveys (9), erratum (6), and articles in press (4).

The time chart of the number of documents published by year (Figure 1), shows that the term gained popularity in scientific literature starting from 2016, and before that was virtually unused by scholars. The first occurrence in the data set is in 2005 (three documents), but until 2016 there were less than 10 documents a year. In 2017 the number of publications suddenly increased to 203, reaching 477 in 2018, 694 in 2019, and 951 in 2020. Compared with the number of documents mentioning “social media”, another steadily growing topic, those mentioning “fake news” were 0.1 percent, on average, between 2010 and 2016, 2.5 percent in 2017, 5.1 percent in 2018, 6.5 percent in 2019, and 7.1 percent in 2020. The growing interest in the phenomenon is also testified by the trend in news articles mentioning the term and by the trend in Google online search (Figure 1).

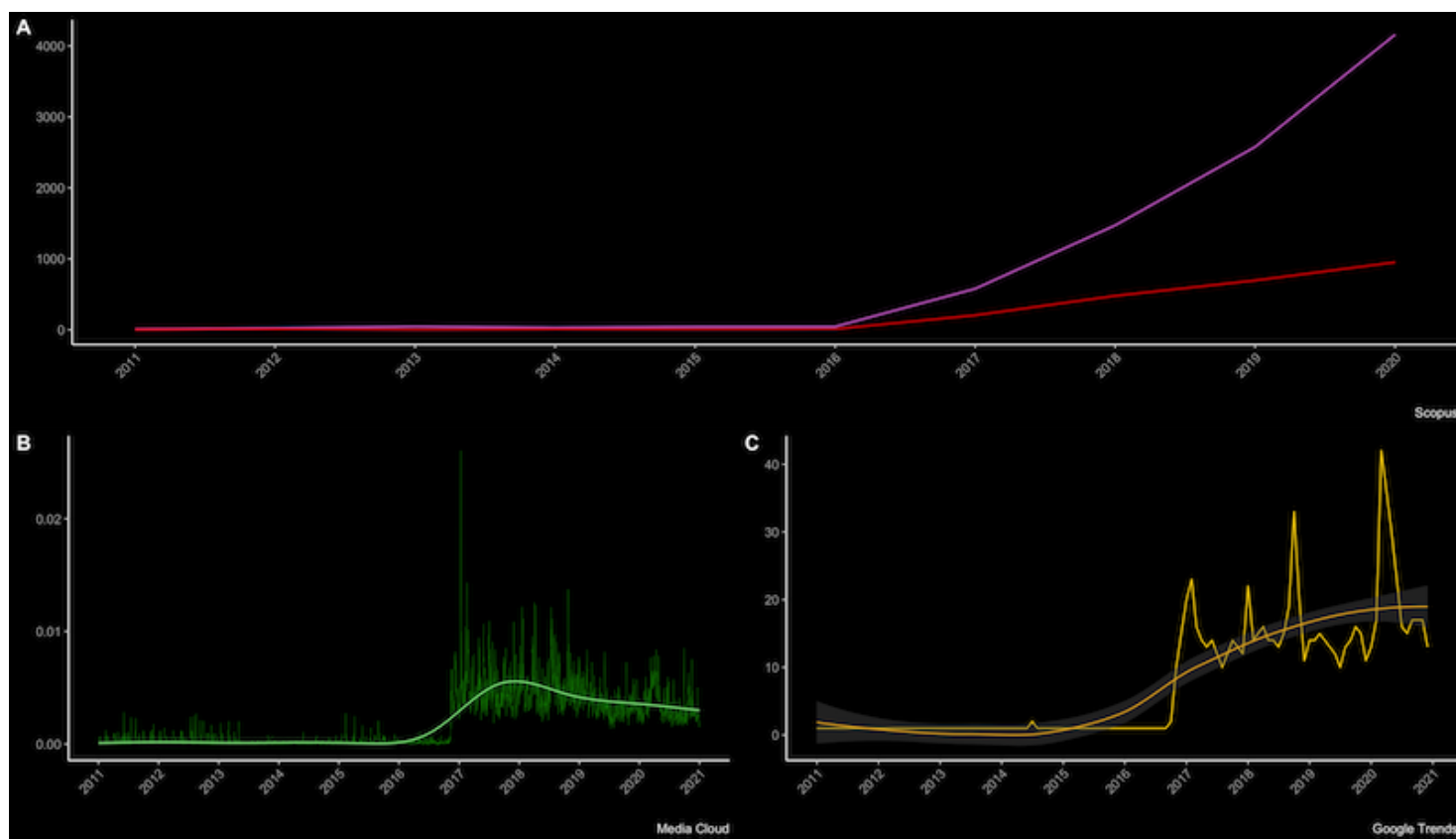


Figure 1: Number of scientific documents mentioning “fake news” in the text (violet) and in the title and abstract (red) (A), number of online newspaper articles mentioning the term “fake news” (B), and Google search trend for the keyword “fake

news” (C).

Geographical distribution

Among the most prolific countries are the United States (698 documents), U.K. (211), India (182), Germany (112), Australia (110), and Italy (99) (Figure 2). A similar, although not totally overlapping, distribution emerges considering the top countries by corresponding author: U.S. (254 documents), U.K. (75), India (56), Australia (51), Germany (45), Spain (39), Korea (38), and Italy (36).

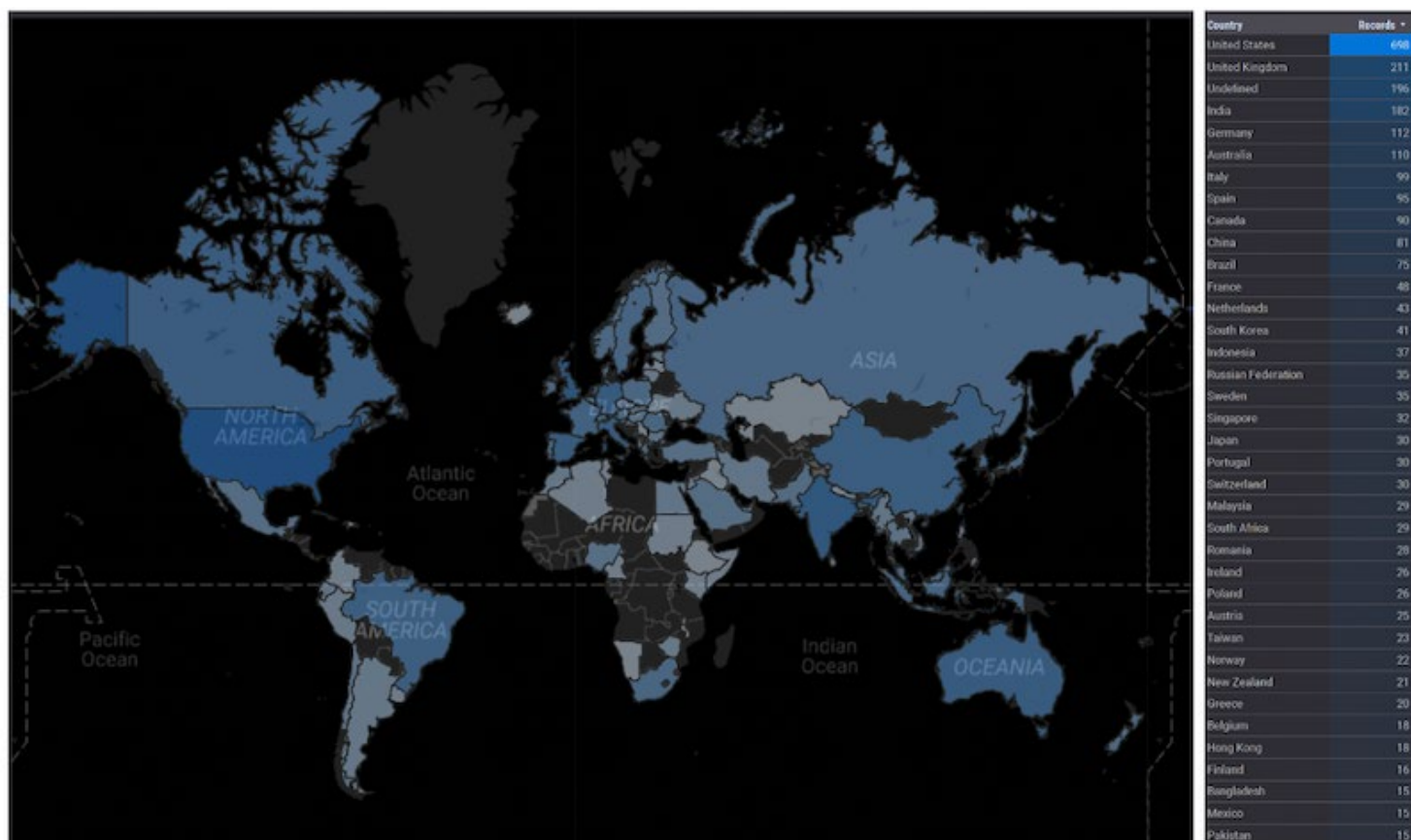


Figure 2: Number (log) of scientific documents mentioning “fake news” in the title or abstract by authors’ affiliation country. The table on the right includes countries with at least 15 associated documents.

These data might suggest a stronger focus on the U.S. and more generally on Western countries, which is reasonable, considering the role of American elections and the former president of the United States, Donald Trump, in popularizing the term “fake news”. However, despite a large number of documents being linked to the U.S. and European countries, figuring among productive countries are also India (182 documents), China (81 documents), South Korea (41 documents), and Indonesia (37 documents), besides others.

Disciplinary perspectives

Fake news is a multifaceted problem and can be tackled from many disciplinary perspectives. The relation between this phenomenon and social media makes it clear why the top discipline by number of contributions is computer science (1,138 documents). Social sciences comes second (939 documents), and among the top ten academic areas are featured scientific, social, and humanistic disciplines such as engineering (346), mathematics (320), arts and

humanities (300), decision science (230), medicine (203), business management and accounting (149), psychology (97), and physics and astronomy (63) (Figure 3).

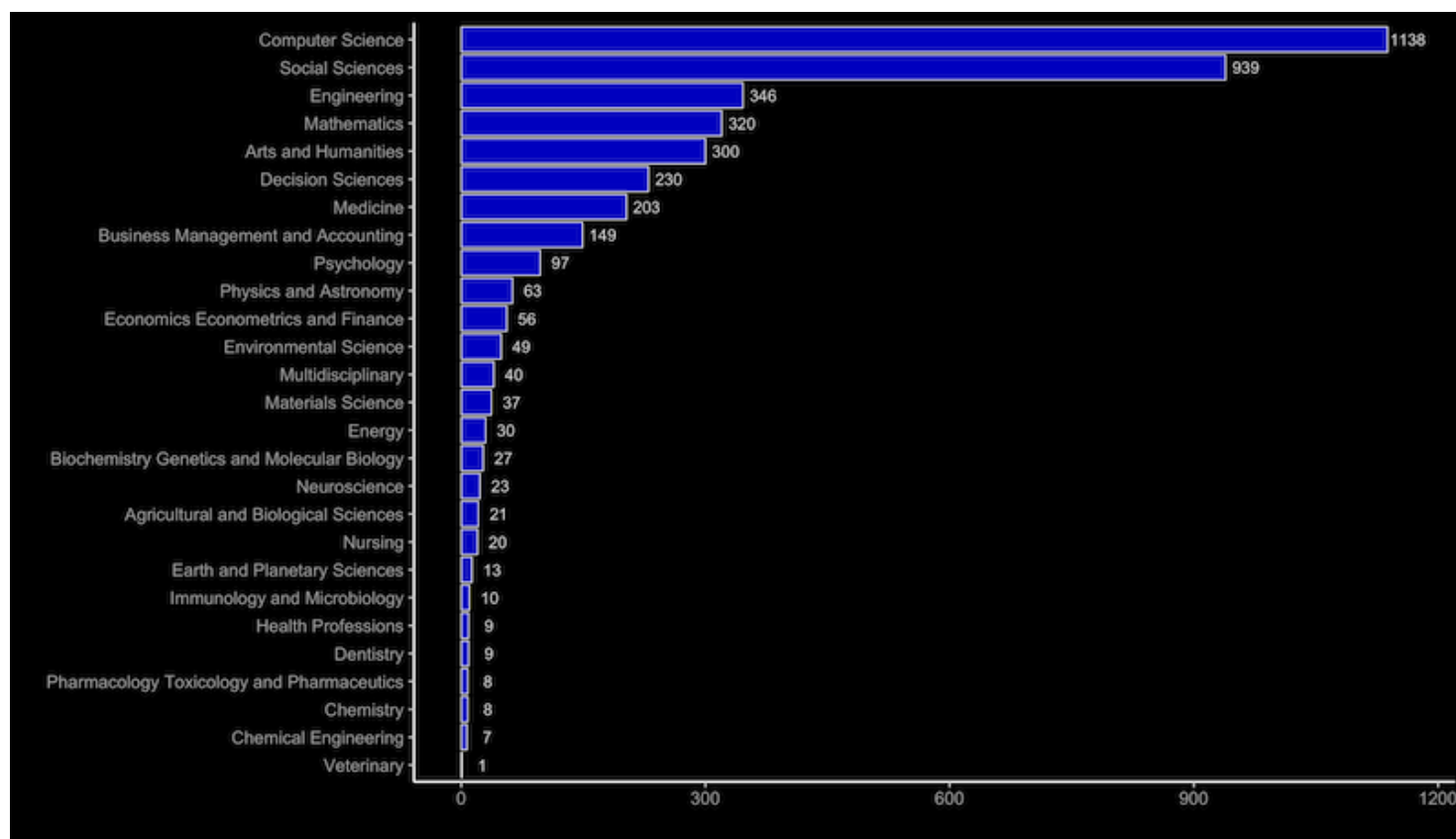


Figure 3: Number of scientific documents mentioning “fake news” in the title or abstract by discipline.

Thematic analysis

By considering the keywords used to describe document topics (Figure 4), besides a focus on social media (including Twitter, with 86 occurrences, and Facebook, with 64 occurrences), what emerges is a substantial methodological interest in fake news detection (with keywords such as machine learning, deep learning, learning algorithms, or artificial intelligence), and computational methods (natural language process, text processing), which are necessary to deal with large amounts of data (big data). This picture is consistent with the high number of contributions published in fields related to computer science.

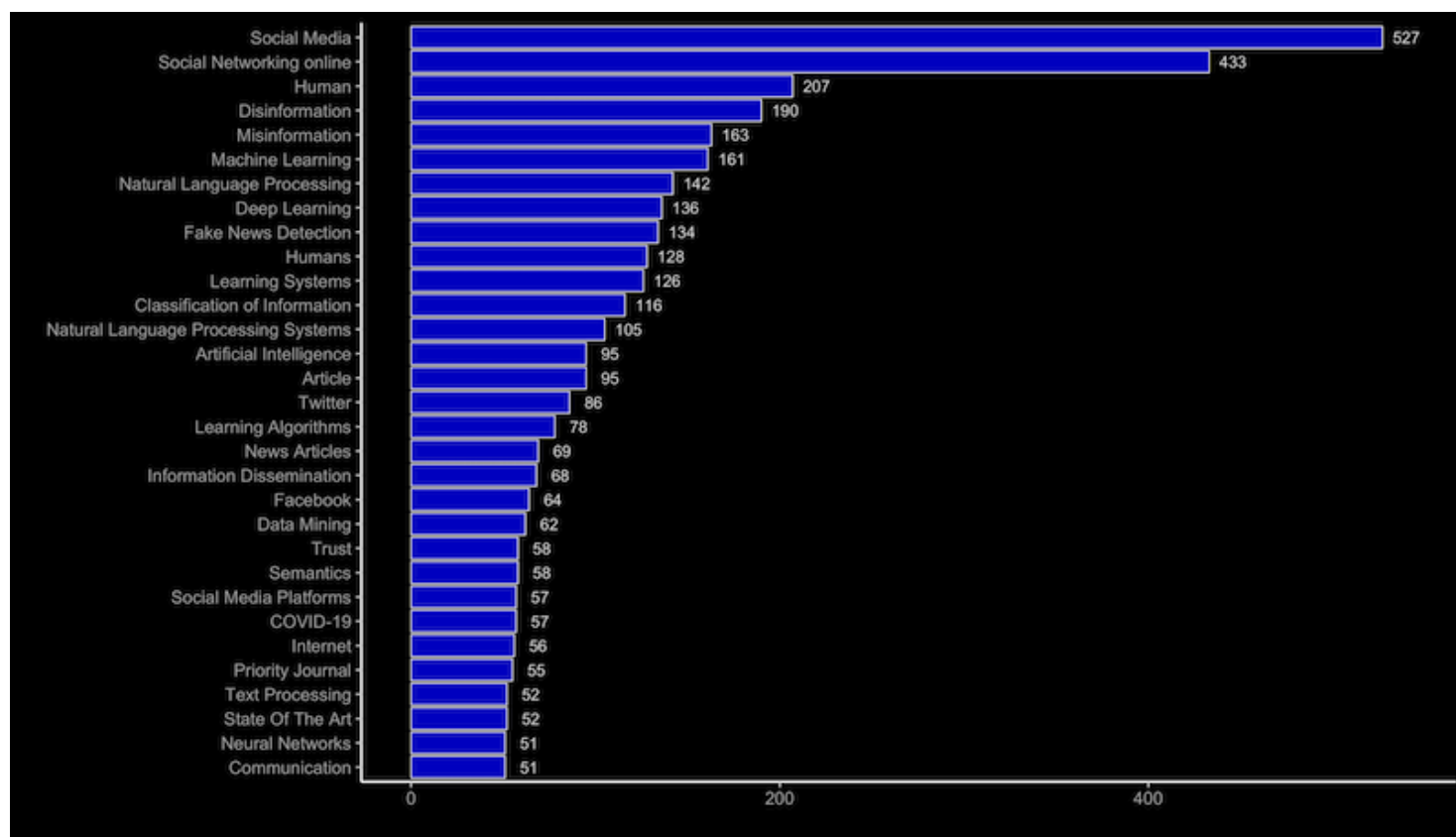


Figure 4: The most used keywords in the investigated set of scientific documents.

Keywords like journalism, information system, communication, and politics, also suggest a particular attention to the socio-political and communicative sides of the problem. Keywords like “pandemic” and “COVID-19” point out the relevance of the misinformation problem during the current epidemiological crisis due to the SARS-CoV-2 virus.

The results of text clustering confirmed and further developed these findings by identifying three main thematic clusters, depicting the field of study of fake news as tripartite in a techno, social, and political dimension (Figure 5). The first cluster (bottom left quadrant) represents the semantics of computational approaches applied to fake news detection. On the top right quadrant there is the cluster representing the social and individual side of fake news, including, for instance, the spread and sharing of misinformation, and the exposure to and the consumption of misinformation, news and information by users. The third cluster, on the bottom right quadrant, represents the public dimension of the problem, with attention, for instance, to journalism, democracy, and the public sphere.

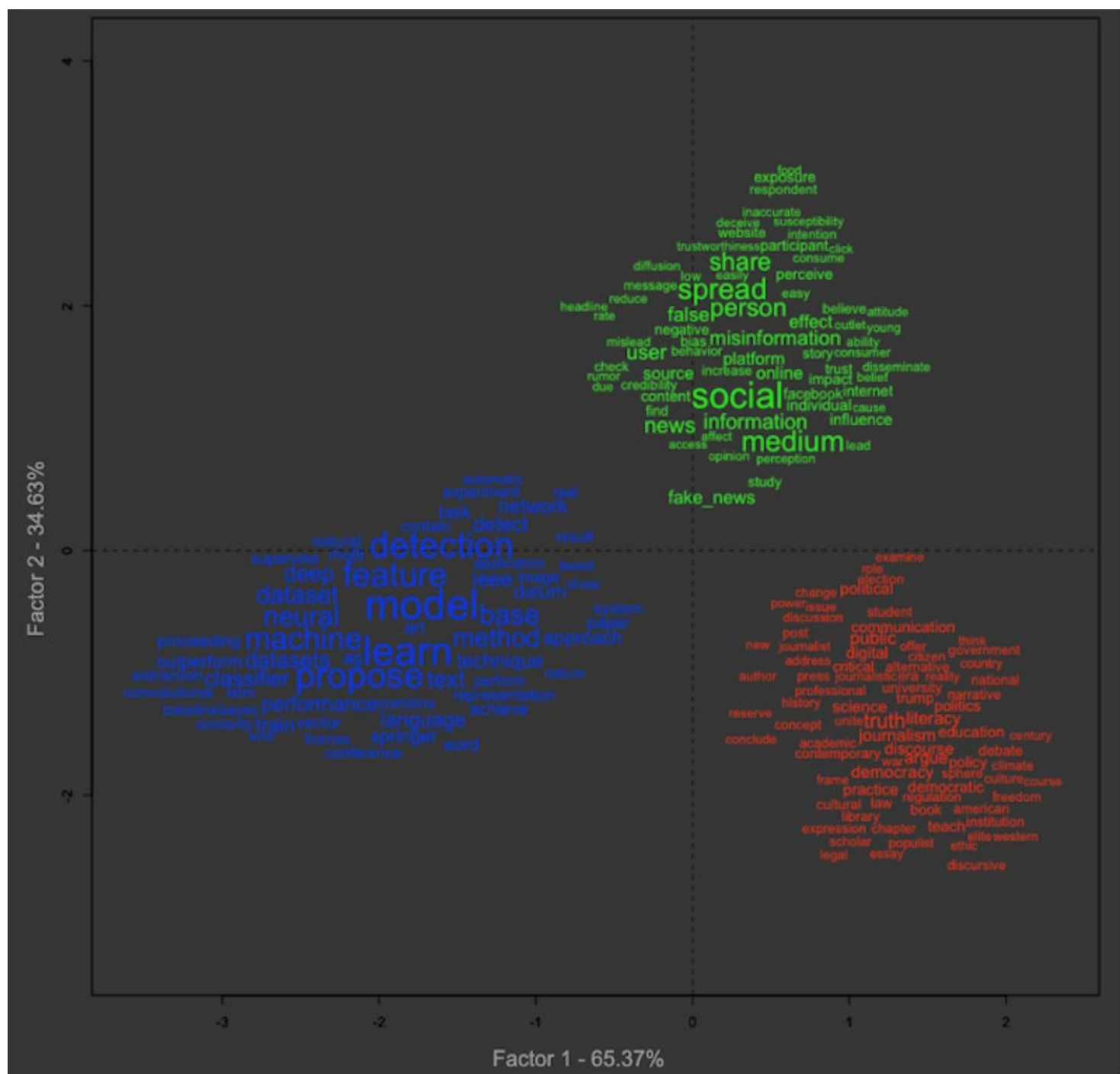


Figure 5: Factor analysis plan following descending hierarchical classification of textual content of titles and abstracts of scientific documents mentioning “fake news” in their title or abstract.

Looking at the estimated prevalence of topics by year resulting from the structural topic model analysis (Figure 6), what emerges is a significant growth over time of interest in certain areas, first and foremost detection methods in 2019 and 2020 (Topic 11), or at the intersection between platforms and user behavior (Topic 15). Instead, topics that were trending and even epitomized the outset of the phenomenon, gradually ceased to attract the scholarly interest. This is the case, for instance, with post-truth (Topic 10), and the presidential elections won by Trump in 2016 (Topic 3). Finally, new topics like COVID-19 and health-misinformation show a rapid growth, and can be expected to further galvanize research in this area (Topic 8).

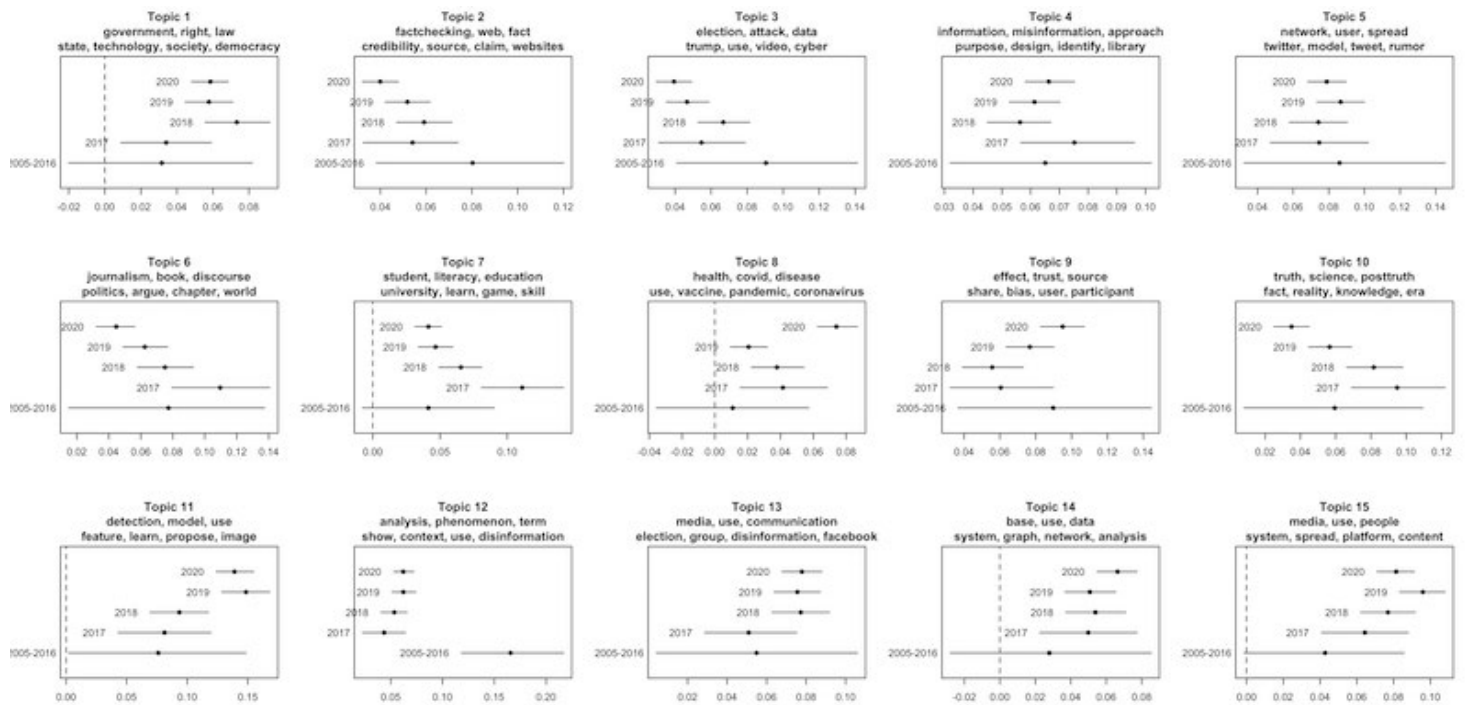


Figure 6: Prevalence of topics by year.

The ten most cited papers on “fake news”

The ten most cited papers have more than 200 citations on Scopus. The paper with the highest number of citations (1,178) is “Social media and fake news in the 2016 election”, by Allcott and Gentzkow, published in the *Journal of Economic Perspectives* (Allcott and Gentzkow, 2017). The paper offers a theoretical and empirical background to the debate surrounding the role of fake news in the election of Donald Trump. It sketches a model of the fake news media market, presents data on fake news consumption and exposure in the run-up to the election, and on the influence of partisanship and other correlates, including social network segregation, on the ability to recognize false news. In the conclusions, the authors observe that “if one fake news article were about as persuasive as one TV campaign ad, the fake news in our database would have changed vote shares by an amount on the order of hundredths of a percentage point. This is much smaller than Trump’s margin of victory in the pivotal states on which the outcome depended” [4].

The second most cited paper (684 citations) is entitled “The science of fake news: Addressing fake news requires a multidisciplinary effort”. The paper, published by Lazer and co-authors appeared in *Science* (Lazer, *et al.*, 2018), provides a definition of fake news, an overview of the related historical context, hints on the prevalence and impact of the problem and on potential interventions. In the section on prevalence and impact, the authors notice that “Evaluations of the medium-to-long-run impact on political behavior of exposure to fake news (for example, whether and how to vote) are essentially nonexistent in the literature. The impact might be small — evidence suggests that efforts by political campaigns to persuade individuals may have limited effects. However, mediation of much fake news via social media might accentuate its effect [...] There exists little evaluation of the impacts of fake news in these regards” [5].

The paper “Defining ‘fake news’: A typology of scholarly definitions”, authored by Edson C. Tandoc, Jr., Zheng Wei Lim, and Richard Ling (2018) and published in *Digital Journalism*, outlines a typology of fake news. Based on a review of 34 academic articles that used the term between 2003 and 2017, they identified six ways that previous studies had defined fake news: news satire, news parody, news fabrication, photo manipulation, advertising and public relations, and propaganda. Satire uses “humor or exaggeration to present audiences with news updates” often by using “the style of a television news broadcast” [6]. Parody “also uses a presentation format which mimics mainstream news media” but differs from satires in its use of non-factual information to inject humor. News satire and parody enable

“critiques of both people in power and also of the news media. [...] Parody news, as well as news satire, are different from other forms of fake news in that there is the assumption that both the author and the reader of the news share the gag” [7]. News fabrication “refers to articles which have no factual basis but are published in the style of news articles to create legitimacy. Unlike parody, there is no implicit understanding between the author and the reader that the item is false. Indeed, the intention is often quite the opposite. The producer of the item often has the intention of misinforming” [8]. Photo manipulation is the “manipulation of real images or videos to create a false narrative” [9]. In the context of advertising and public relations, “fake news was defined as ‘when public relations practitioners adopt the practices and/or appearance of journalists in order to insert marketing or other persuasive messages into news media’” [10]. Finally, propaganda “refers to news stories which are created by a political entity to influence public perceptions. The overt purpose is to benefit a public figure, organization or government” [11].

The conference paper by Gupta, Lamba, Kumaraguru, and Joshi (2013), entitled “Faking Sandy: Characterizing and identifying fake images on Twitter during Hurricane Sandy”, published in the *Proceedings of the 22nd International Conference on World Wide Web*, focuses on the spread of fake images on Twitter during Hurricane Sandy (2012), a case described by Tandoc, *et al.* (2018) as “photo manipulation”. They apply and analyze the effectiveness of machine learning classifiers to distinguish fake from real images.

The paper “The Daily Show: Discursive integration and the reinvention of political journalism”, authored by Baym (2005) and published in *Political Communication*, focuses on the “fake news” satirical program “The Daily Show” with Jon Stewart: “Unquestionably, its primary approach is comedy, and much of the show’s content is light and, at times, vacuous. Often, however, the silly is interwoven with the serious, resulting in an innovative and potentially powerful form of public information. [...] Lying just beneath or perhaps imbricated within the laughter is a quite serious demand for fact, accountability, and reason in political discourse” [12].

The paper entitled “Beyond misinformation: Understanding and coping with the ‘post-truth’ era”, authored by Lewandowsky, Ecker, and Cook (2017) and published in the *Journal of Applied Research in Memory and Cognition*, is an exploration of the misinformation landscape with particular attention to “post-truth”, a term that gained popularity in 2016 along with “fake news”. Post-truth was defined by *Oxford Dictionaries* as “relating to or denoting circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief” [13], and can be framed as an “alternative epistemology that does not conform to conventional standards of evidentiary support” created by “political drivers” [14]. After discussing some trends that may have contributed to the emergence of a post-truth world over the last few decades — namely processes related to the decline in social capital and shifting values, growing inequality and increasing polarization, declining trust in science, politically asymmetric credulity, and the evolution of the media landscape — the authors characterize the post-truth discourse and politics [15], concluding that “post-truth misinformation has arguably been designed and used as a smokescreen to divert attention from strategic political actions or challenges” and the “resolution requires political mobilization and public activism” [16]. Also in the light of the political and socio-cultural trends that made it possible, the misinformation of the post-truth world can no longer be considered “an isolated failure of individual cognition that can be corrected with appropriate communication tools” [17].

The book *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media* by Gillespie (2018), is an investigation into platform politics regarding content moderation: “The hard questions being asked now, about freedom of expression and virulent misogyny and trolling and breastfeeding and pro-anorexia and terrorism and fake news, are all part of a fundamental reconsideration of social media platforms. [...] If moderation should not be conducted the way it has, what should take its place?” [18].

The conference paper by Conroy, Rubin, and Chen (2015), entitled “Automatic deception detection: Methods for finding fake news”, and published in the *Proceedings of the Association for Information Science and Technology*, provides a survey of “the current state-of-the-art technologies that are instrumental in the adoption and development of fake news detection” and “a typology of several varieties of veracity assessment methods emerging from two major categories — linguistic cue approaches (with machine learning), and network analysis approaches”, and proposes “an innovative hybrid approach that combines linguistic cue and machine learning, with network-based behavioral data” [19].

The conference paper “CSI: A hybrid deep model for fake news detection”, authored by Ruchansky, Seo, and Liu (2017) and published in the *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, proposes another approach to fake news detection, named CSI (Capture, Score, and Integrate).

The last paper in the top ten is also the most recent one. It is entitled “Fake news on Twitter during the 2016 US

presidential election”, and was published in *Science* by Grinberg, Joseph, Friedland, Swire-Thompson, and Lazer (2019). The paper examines the exposure to, and the sharing of fake news on Twitter by voters, finding that “only 1% of individuals accounted for 80% of fake news source exposures, and 0.1% accounted for nearly 80% of fake news sources shared”, and that the individuals “most likely to engage with fake news sources were conservative leaning, older, and highly engaged with political news” [20]: “Although 6% of people who shared URLs with political content shared content from fake news sources, the vast majority of fake news shares and exposures were attributable to tiny fractions of the population. [...] For the average panel member, content from fake news sources constituted only 1.18% of political exposures, or about 10 URLs during the last month of the election campaign. [...] we found that the vast majority of political exposures, across all political groups, still came from popular non-fake news sources. This is reassuring in contrast to claims of political echo chambers and fake news garnering more engagement than real news during the election” [21].

Table 1: Top 10 most cited documents.						
Year	Document title	Authors	Publication title	Volume	Number	Citations (Scopus)
2017	Social media and fake news in the 2016 election	Allcott H., Gentzkow M.	<i>Journal of Economic Perspectives</i>	31	2	1,178
2018	The science of fake news: Addressing fake news requires a multidisciplinary effort	Lazer D.M.J., Baum M.A., Benkler Y., Berinsky A.J., Greenhill K.M., Menczer F., Metzger M.J., Nyhan B., Pennycook G., Rothschild D., Schudson M., Sloman S.A., Sunstein C.R., Thorson E.A., Watts D.J., Zittrain J.L.	<i>Science</i>	359	6380	684
2018	Defining “fake news”: A typology of scholarly definitions	Tandoc E.C., Lim Z.W., Ling R.	<i>Digital Journalism</i>	6	2	355
2013	Faking Sandy: Characterizing and identifying fake images on Twitter during Hurricane Sandy	Gupta A., Lamba H., Kumaraguru P., Joshi A.	<i>WWW 2013 Companion: Proceedings of the 22nd International Conference on World Wide Web</i>			273
2005	The Daily Show: Discursive integration and	Baym G.	<i>Political</i>	22	3	272

	the reinvention of political journalism		<i>Communication</i>			
2017	Beyond misinformation: Understanding and coping with the “post-truth” era	Lewandowsky S., Ecker U.K.H., Cook J.	<i>Journal of Applied Research in Memory and Cognition</i>	6	4	253
2018	<i>Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media</i>	Gillespie T.				250
2015	Automatic deception detection: Methods for finding fake news	Conroy N.J., Rubin V.L., Chen Y.	<i>Proceedings of the Association for Information Science and Technology</i>	52	1	249
2017	CSI: A hybrid deep model for fake news detection	Ruchansky N., Seo S., Liu Y.	<i>International Conference on Information and Knowledge Management, Proceedings</i>			225
2019	Political science: Fake news on Twitter during the 2016 U.S. presidential election	Grinberg N., Joseph K., Friedland L., Swire-Thompson B., Lazer D.	<i>Science</i>	363	6425	201

Discussion and conclusions

Research on “fake news” started during and in the aftermath of the 2016 U.S. presidential election and shows no sign of stopping. Up to 2020, 9,015 scientific documents mentioned the term “fake news” somewhere in their full text, and 2,368 documents, written by 5,060 authors and published on 1,225 different sources, mentioned it in the title or abstract.


The historical origin of the problem can help to interpret the predominant U.S. and Western focus, as suggested by the analysis of the authors’ affiliations, but non-Western countries’ contributions are present as well. Analyses that take into account different cultures, societies, and political regimes can certainly help to better understand a worldwide issue like fake news, and possibly, to develop some solutions.

The analysis of the disciplinary areas showed a marked methodological interest in fake news detection by means of machine learning algorithms. Not only has this topic attracted special attention in the last couple of years, but papers

on fake news detection are also among the top ten most cited papers on Scopus (Conroy, *et al.*, 2015; Gupta, *et al.*, 2013; Ruchansky, *et al.*, 2017). Of course, there are good reasons for the interest in detection, since it would be necessary to combat fake news on social media platforms. At the same time, a social science approach can unpack the phenomenon from a qualitative and theoretical point of view.

Even though the focus on the social, individual, public and political dimension of the phenomenon comes only second, in terms of quantity of publications, there are also extensive analyses on the societal and technological trends related to the fake news problem (Gillespie, 2018; Lewandowsky, *et al.*, 2017), along with research on the actual exposure and potential impact of fake news on citizens and political behavior (Allcott and Gentzkow, 2017; Grinbe, *et al.*, 2019), conceptual and terminological analyses (Tandoc, *et al.*, 2018), and general introductions to the problem (Lazer, *et al.*, 2018).

In general, this research showed that the expression “fake news”, despite receiving critical attention for being bandied about as a blanket term and failing to describe the complex landscape of online misinformation, not only keeps being widely used, but continues to be endorsed by a growing trend in publications resorting to it. The current COVID-19 pandemic, which appears among the most used keywords despite being a recent phenomenon, has probably given a strong impulse to research on the topic.

A constantly growing field of study like the one concerning fake news requires scholars to have a general overview of the scientific productions on the topic, and a systematic, quantitative literature review like the one carried out in these pages can be of help. The variety of perspectives and topics addressed by scholars, however, also means that future analyses will need to focus on more specific themes and include more variables (Chan and Grill, 2020). To further develop analysis of the field, a line of inquiry might combine quantitative and qualitative approaches to trace and describe — within the broad trend of attention to fake news and mis/disinformation — the sub-trends that have typified the development of the field. The time series analysis of the topics revealed, indeed, that scholarly attention shifted from ephemeral and rapidly sidelined trends such as, to cite an example, post-truth, to more structural lines of inquiry, such as fake news detection. It also showed the impact of real-world events on academic studies, starting from events like the 2016 presidential elections and the COVID-19 pandemic. Identifying and distinguishing between these strands of research and shedding light on their determinants, analyzing them in depth through a qualitative approach, could provide a fine-grained description of the field and allow scholars to take a step back to critical reflect on its construction, thus favoring its future development. 

About the author

Nicola Righetti is a researcher in computational communication science at the Department of Communication of the University of Vienna.

E-mail: nicola [dot] righetti [at] univie [dot] ac [dot] at

Notes

- [1.](#) Lazer, *et al.*, 2018, p. 1,094.
- [2.](#) Allcott and Gentzkow, 2017, p. 213.
- [3.](#) Lazer, *et al.*, 2018, p. 1,094.
- [4.](#) Allcott and Gentzkow, 2017, p. 232.
- [5.](#) Lazer, *et al.*, 2018, p. 1,095.
- [6.](#) Tandoc, *et al.*, 2018, p. 141.
- [7.](#) Tandoc, *et al.*, 2018, p. 142.
- [8.](#) Tandoc, *et al.*, 2018, p. 143.

9. Tandoc, *et al.*, 2018, p. 144.
10. Tandoc, *et al.*, 2018, p. 145.
11. Tandoc, *et al.*, 2018, p. 146.
12. Baym, 2005, p. 273.
13. Oxford Languages, “Word of the year 2016,” at <https://languages.oup.com/word-of-the-year/2016/>.
14. Lewandowsky, *et al.*, 2017, p. 356.
15. Lewandowsky, *et al.*, 2017, pp. 357–362.
16. Lewandowsky, *et al.*, 2017, pp. 364–365.
17. Lewandowsky, *et al.*, 2017, p. 353.
18. Gillespie, 2018, pp. 12–13.
19. Conroy, *et al.*, 2015, p. 1.
20. Grinberg, *et al.*, 2019, p. 374.
21. Grinberg, *et al.*, 2019, p. 378.

References

- Hunt Allcott and Matthew Gentzkow, 2017. “Social media and fake news in the 2016 election,” *Journal of Economic Perspectives*, volume 31, number 2, pp. 211–236.
doi: <https://doi.org/10.1257/jep.31.2.211>, accessed 20 May 2021.
- Geoffrey Baym, 2005. “The Daily Show: Discursive integration and the reinvention of political journalism,” *Political Communication*, volume 22, number 3, pp. 259–276.
doi: <https://doi.org/10.1080/10584600591006492>, accessed 20 May 2021.
- Chung-hong Chan and Christiane Grill, 2020. “The highs in communication research: Research topics with high supply, high popularity, and high prestige in high-impact journals,” *Communication Research* (30 July).
doi: <https://doi.org/10.1177/0093650220944790>, accessed 20 May 2021.
- Nadia K. Conroy, Victoria L. Rubin, and Yimin Chen, 2015. “Automatic deception detection: Methods for finding fake news,” *Proceedings of the Association for Information Science and Technology*, volume 52, number 1, pp. 1–4.
doi: <https://doi.org/10.1002/prat.2015.145052010082>, accessed 20 May 2021.
- Edwin S. Dalmaijer, Joram Van Rheede, Edwin V. Sperr, and Juliane Tkotz, 2021. “Banana for scale: Gauging trends in academic interest by normalising publication rates to common and innocuous keywords,” *arXiv:2102.06418* (12 February), at <https://arxiv.org/abs/2102.06418>, accessed 20 May 2021.
- Jana Laura Egelhofer and Sophie Lecheler, 2019. “Fake news as a two-dimensional phenomenon: A framework and research agenda,” *Annals of the International Communication Association*, volume 43, number 2, pp. 97–116.
doi: <https://doi.org/10.1080/23808985.2019.1602782>, accessed 20 May 2021.
- Jana Laura Egelhofer, Loes Aaldering, Jakob-Moritz Eberl, Sebastian Galyga, and Sophie Lecheler, 2020. “From novelty to normalization? How journalists use the term ‘fake news’ in their reporting,” *Journalism Studies*, volume 21, number 10, pp. 1,323–1,343.
doi: <https://doi.org/10.1080/1461670X.2020.1745667>, accessed 20 May 2021.
- Johan Farkas and Jannick Schou, 2018. “Fake news as a floating signifier: Hegemony, antagonism and the politics of falsehood,” *Javnost — The Public*, volume 25, number 3, pp. 298–314.

doi: <https://doi.org/10.1080/13183222.2018.1463047>, accessed 20 May 2021.

Fabio Giglietto, Laura Iannelli, Augusto Valeriani, and Luca Rossi, 2019. “‘Fake news’ is the invention of a liar: How false information circulates within the hybrid news system,” *Current Sociology*, volume 67, number 4, pp. 625–642. doi: <https://doi.org/10.1177/0011392119837536>, accessed 20 May 2021.

Tarleton Gillespie, 2018. *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. New Haven, Conn.: Yale University Press.

Nir Grinberg, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson, and David Lazer, 2019. “Fake news on Twitter during the 2016 U.S. presidential election,” *Science*, volume 363, number 6425 (25 January), pp. 374–378. doi: <https://doi.org/10.1126/science.aau2706>, accessed 20 May 2021.

Aditi Gupta, Hemank Lamba, Ponnurangam Kumaraguru, and Anupam Joshi, 2013. “Faking Sandy: characterizing and identifying fake images on twitter during Hurricane Sandy,” *WWW '13: Proceedings of the 22nd International Conference on World Wide Web*, pp. 729–736. doi: <https://doi.org/10.1145/2487788.2488033>, accessed 20 May 2021.

Caroline Jack, 2017. “Lexicon of lies: Terms for problematic information,” *Data & Society*, at https://datasociety.net/pubs/oh/DataAndSociety_LexiconofLies.pdf, accessed 12 March 2021.

David M.J. Lazer, Matthew A. Baum, Yoichai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts, and Jonathan L. Zittrain, 2018. “The science of fake news,” *Science*, volume 359, number 6380 (9 March), pp. 1,094–1,096. doi: <https://doi.org/10.1126/science.aao2998>, accessed 20 May 2021.

Stephan Lewandowsky, Ullrich K.H. Ecker, and John Cook, 2017. “Beyond misinformation: Understanding and coping with the ‘post-truth’ era,” *Journal of Applied Research in Memory and Cognition*, volume 6, number 4, pp. 353–369. doi: <https://doi.org/10.1016/j.jarmac.2017.07.008>, accessed 20 May 2021.

Alice Marwick and Rebecca Lewis, 2017. “Media manipulation and disinformation online,” *Data & Society* (15 May), at <https://datasociety.net/library/media-manipulation-and-disinfo-online/>, accessed 20 May 2021.

Pierre Ratinaud, 2008. “IRAMUTEQ (Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires),” at <http://www.iramuteq.org>, accessed 12 March 2021.

Max Reinert, 1990. “Alceste une méthodologie d’analyse des données textuelles et une application: Aurelia De Gerard De Nerval,” *Bulletin of Sociological Methodology/Bulletin de méthodologie sociologique*, volume 26, number 1, pp. 24–54. doi: <https://doi.org/10.1177/075910639002600103>, accessed 20 May 2021.

Margaret E. Roberts, Brandon M. Stewart, and Dustin Tingley, 2019. “Stm: An R package for structural topic models,” *Journal of Statistical Software*, volume 91, number 2, pp. 1–40. doi: <https://doi.org/10.18637/jss.v091.i02>, accessed 20 May 2021.

Natali Ruchansky, Sungyong Seo, and Yan Liu, 2017. “CSI: A hybrid deep model for fake news detection,” *CIKM '17: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pp. 797–806. doi: <https://doi.org/10.1145/3132847.3132877>, accessed 20 May 2021.

Scopus, 2020. “Scopus content coverage guide,” at https://www.elsevier.com/_data/assets/pdf_file/0007/69451/Scopus_ContentCoverage_Guide_WEB.pdf, accessed 12 March 2021.

Edson C. Tandoc, Jr., Zheng Wei Lim, and Richard Ling, 2018. “Defining ‘fake news’: A typology of scholarly definitions,” *Digital Journalism*, volume 6, number 2, pp. 137–153. doi: <https://doi.org/10.1080/21670811.2017.1360143>, accessed 20 May 2021.

TaSC, 2020. “The media manipulation casebook,” at <https://mediamanipulation.org>, accessed 12 March 2021.

Tommaso Venturini, 2019. "From fake to junk news: The data politics of online virality," In: Didier Bigo, Engin Isin, and Evelyn Ruppert (editors). *Data politics: Worlds, subjects, rights*. London: Routledge.
doi: <https://doi.org/10.4324/9781315167305>, accessed 20 May 2021.

Claire Wardle, 2017. "Fake news. It's complicated," *First Draft* (16 February), at <https://firstdraftnews.org/latest/fake-news-complicated/>, accessed 12 March 2021.

Editorial history

Received 15 March 2021; revised 4 May 2021; accepted 20 May 2021.



This paper is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

Four years of fake news: A quantitative analysis of the scientific literature
by Nicola Righetti.

First Monday, volume 26, number 6 (June 2021).

doi: <http://dx.doi.org/10.5210/fm.v26i6.11645>