Article

American Behavioral Scientist
2021, Vol. 65(2) 180–212

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0002764219878224
journals.sagepub.com/home/abs
SAGE

"Fake News" Is Not Simply False Information: A Concept Explication and Taxonomy of Online Content

Maria D. Molina[1], S. Shyam Sundar[1], Thai Le[1], and Dongwon Lee[1]

Abstract

As the scourge of "fake news" continues to plague our information environment, attention has turned toward devising automated solutions for detecting problematic online content. But, in order to build reliable algorithms for flagging "fake news," we will need to go beyond broad definitions of the concept and identify distinguishing features that are specific enough for machine learning. With this objective in mind, we conducted an explication of "fake news" that, as a concept, has ballooned to include more than simply false information, with partisans weaponizing it to cast aspersions on the veracity of claims made by those who are politically opposed to them. We identify seven different types of online content under the label of "fake news" (false news, polarized content, satire, misreporting, commentary, persuasive information, and citizen journalism) and contrast them with "real news" by introducing a taxonomy of operational indicators in four domains—message, source, structure, and network—that together can help disambiguate the nature of online news content.

Keywords

fake news, false information, misinformation, persuasive information

"Fake news," or fabricated information that is patently false, has become a major phenomenon in the context of Internet-based media. It has received serious attention in a variety of fields, with scholars investigating the antecedents, characteristics, and consequences of its creation and dissemination. Some are primarily interested in the

[1]Penn State University, University Park, PA, USA

Corresponding Author:
Maria D. Molina, Media Effects Research Laboratory, Donald P. Bellisario College of Communications, Penn State University, 115 Carnegie Building, University Park, PA 16802, USA.
Email: mdm63@psu.edu

nature of misinformation contained in false news, so that we can better detect it and distinguish it from real news. Others focus on the susceptibility of users—why we fall for false news and how we can protect ourselves from this vulnerability. Both are geared toward improving media literacy to protect consumers from false information.

Websites like Snopes and Politifact explicitly address the issue by verifying information in the news cycle with the help of an army of human fact-checkers. However, human fact-checking can be time consuming and subject to human foibles such as subjectivity and being limited by prior experiences (Vorhies, 2017). An alternative that has been proposed is the use of machine algorithms to facilitate the detection of false information (Conroy, Rubin, & Chen, 2015; Wang, 2017). Given the enormity of the fabricated news problem, machine-based solutions seem inevitable for tackling the scope and speed with which it is created and disseminated, especially around the time of elections, disasters, crises, and other developing stories. However, in order to develop reliable algorithms for detecting fabricated news, we have to be very disciplined in defining it and differentiating it from legitimate news.

Several scholars have led the way in defining and classifying "fake news." Southwell, Thorson, and Sheble (2017) pinpoint conceptual differences between misinformation and disinformation, and discuss in detail the grounding of truth necessary for defining misinformation. Lazer et al. (2018) reminds us of the importance of process and intent when conceptualizing "fake news," while Jack (2017) further disentangles the conceptual differences and similarities among an array of concepts associated with problematic information, ranging from disinformation to propaganda. Tandoc, Lim, and Ling (2018), on the other hand, analyze how the term "fake news" specifically has been used by scholars, and developed a typology based on facticity and intention to deceive. They proceeded to place different kinds of online content on these two dimensions, with Propaganda, for example, scoring high on both, whereas *fabrication* is low on facticity but high on deceptive intention and *satire* being high on facticity and low on intention to deceive. Such classification of different types of "fake news" is a useful starting point in enhancing our understanding of the phenomenon. However, we need more such distinguishing characteristics and dimensions, especially those that can be usefully incorporated in automated detection algorithms. Facticity is useful for fact-checking news stories, but cannot be relied on in the case of breaking news about emergent events where no previous information is available. The intent to deceive may probably be inferred by knowing the pedigree of the news source, but can be difficult to establish in a dispositive manner. Thus, for machine learning purposes, we need a more comprehensive definition that can not only distinguish between a variety of "fake news" but also lend itself to operationalization at a very granular level for machine detection.

Previous research has provided conceptual differences and similarities between the many terms associated with "fake news." The purpose of the current research is to go one step further by identifying specific operational features or indicators that can be fed into a machine-learning algorithm to reliably differentiate between different types of content that are associated with the broad label of "fake news." With supervised machine learning, the goal is to develop "algorithms that reason from externally supplied label training instances to produce general hypotheses, which then make predictions about future instances" (Kotsiantis, 2007, p. 249). Through this process, the machine is trained with preestablished indicators, or features, from a training set, to then predict its corresponding category for unseen instances by a set of learned rules. For example, Decision Tree algorithm learns a sequence of yes/no questions conditioned on input features to make classification, or Naïve Bayes family of algorithms models the conditional probability of different "fake news" categories given input features to make predictions. In this process, the best fitting algorithm is selected and its performance metric (e.g., Accuracy, F1 score) is calculated (Murphy, 2012). Moreover, these operational features are interpretable, which is more beneficial than latent features automatically learned by powerful deep learning models that are usually obscure. Classification works through a set of indicators and results in a set of rules (e.g., Decision Tree), meaning that no one rule serves as ground truth for a classification, and overlapping indicators can exist between categories. Importantly, indicators need to be operationalized to a sufficiently concrete degree, a process called features engineering, in order to allow the algorithm to perform its prediction. The goal of this research, thus, is to identify novel indicators or rules that can be used by a machine-learning algorithm to predict if a piece of information is indeed one among the types of content that we label as "fake news." Once they are successfully engineered into algorithms, these indicators can be used in literacy programs to help online users be more attentive to their information environment, providing specific tips that could guide them in identifying if a piece of online content is indeed legitimate news or "fake news."

A complicating factor in this exercise is that "fake news" no longer refers simply to false information. As Vosoughi, Roy, and Aral (2018) point out, the term "fake news" has been "irredeemably polarized" in that it has been co-opted by politicians to refer to any information put out by sources that do not support their partisan positions. Waisbord (2018) calls it a "trope used by right-wing politicians, commentators and activists to castigate critical news organizations" (p. 1867). The central goal of these partisan efforts is to cast aspersions on the veracity of the content by suggesting that it is false. In many cases, such content is not even news where truth or falsehood is relevant, but a piece of commentary that expresses a particular point of view or an incomplete report of an event by a citizen journalist than can be interpreted in multiple ways. By extending the notion of falsehood to these other categories of non-news content, the term "fake news" has been widely weaponized. But, it has also resulted in the application of the label to a confusing array of content that lies at the intersection of legitimately real and patently false information, thereby posing considerable challenges to machine-based classification of "fake news."

To reduce the semantic clutter around the term "fake news" and derive meaningful indicators of the various types of content that has now become part of it, we launched a concept explication (Chaffee, 1991) that uncovers the different theoretical and operational definitions of "fake news" and its associated terms (e.g., misinformation, disinformation, alternative facts, and false news) found in academic research, media articles, trade journals, and other relevant sources. We reviewed publications found through Google Scholar and a library database using the aforementioned keyword searches. We then followed a snowball approach to identify additional relevant articles until we reached a saturation point in terms of variety in theoretical definitions (Saunders et al., 2018).

Through this meaning analysis, a taxonomy of online content was developed with two primary objectives. Our first objective was to pinpoint the key defining characteristics of false information. Knowing the main ingredients of false news will facilitate the development of an algorithm for detection of this type of content. Second, we wanted to identify other types of content that is often confused with false news but is not false news. A conceptual understanding of these types of content will help us better distinguish them from false news and rule them out for machine-detection purposes.

The goal of this article, therefore, lies not only in describing different types of information and misinformation that people encounter online but also in outlining specific similarities and differences between these at the operational level. We do this by first describing the different definitions of false news encountered through our meaning analysis and the current inconsistencies in the basic assumptions of what is and is not false news. We will then propose a theoretical definition of false information and situate it in a new taxonomy of online content, highlighting the characteristics that help distinguish one type of content from another. We compiled these characteristics from the literature reviewed for this concept explication and categorized them into features of content, source, structure, and network in an effort to provide guidance to algorithm designers about the appropriate loci for each type of content. Finally, from the developed taxonomy and characteristics, we derive specific features or indicators for use in a machine-learning algorithm.

Evolution of Fake or Fabricated News: Theoretical and Operational Definitions

Although the interest in "fake news" spiked after the 2016 Presidential election, it is not a new phenomenon. The concept, known as "disinformation" during the World Wars and as "freak journalism" or "yellow journalism" during the Spanish war, can be traced back to 1896 (Campbell, 2001; Crain, 2017). Yellow journalism was also known for publishing content with no evidence and therefore factually incorrect, often for business purposes (Samuel, 2016). In Yarros' (1922) critique of yellow journalism, he characterizes it as "brazen and vicious 'faking,' and reckless disregard of decency, proportion and taste for the sake of increased profits" (p. 410).

As if history were repeating itself, the phenomenon regained attention during the 2016 U.S. Presidential elections. However, what makes fabricated news unique is the information environment we currently live in, where social media are key to dissemination of information and we no longer receive information solely from traditional gatekeepers. Nowadays, it is not necessary to be a journalist and work for a publication to create and disseminate content online. Laypersons write, curate, and disseminate information via online media. Studies show that they may even be preferred over traditional professional sources (Sundar & Nass, 2001). This is particularly troublesome given that individuals find information that agrees with prior beliefs as more credible and reliable, creating an environment that exacerbates misinformation because credible information appears alongside personal opinions (Bode & Vraga, 2015).

Defining "Fake News"

Although we now have a seemingly simple dictionary definition of "fake news" as "false stories that appear to be news, spread on the Internet or using other media, usually created to influence political views or as a joke" (Fake News, 2018), determining what is and what is not false is rather complex. There is considerable disagreement when it comes to determining which content should be considered "fake news" and which should be excluded. This holds especially true as the term "fake news" has become highly political and is often used as a buzzword not only used to describe fabricated information but to undermine the credibility of news organizations or argue against commentary that disagrees with our own opinion (Nielsen & Graves, 2017; Tandoc et al., 2018). Moreover, classifying a piece of content as false requires a grounding of a universal truth, which can be a difficult endeavor that requires collective consensus (see Southwell et al., 2017). We acknowledge the complexities and ambiguity of the term. As such, we will use the term "fake news" to provide a historical background and explain the current uses of the term. We also use it as a broad umbrella term with seven categories under it, of which false information (of false news) is one of them. This is because we are interested in identifying news that is fabricated or fictitious and disseminated online. As such, we need to contain the concept so that we focus on identifying this type of problematic information instead of opening a debate that would undermine the goal of our project. We also use the umbrella term "fake news" in our taxonomy rather than the label of misinformation as suggested by scholars in the field (i.e., Southwell et al., 2017) because misinformation is by definition false information determined based on a grounding of truth and applies only to informationally oriented content. However, content that is often accused of being "fake news" is not always intended for informational purposes. For example, some of it is clearly persuasive content that is intended to persuade, not necessarily inform. Satirical accounts of news events (the original meaning of "fake news") is also not meant for information, but for entertainment. Then, there are other kinds of content that although are not "real news" per se, are not falsehoods or erroneous either, for example, commentary or citizen journalism. In other words, our taxonomy goes beyond identifying false information (whether intentional or not), to identifying the different types of content available online that might be mislabeled as false or "fake." We acknowledge that using the term "fake news" is contentious because of its highly politicized nature. Nevertheless, the goal of our project is to precisely illustrate how the term "fake news" has been used to refer to different types of content online, that in reality are different form each other, theoretically and empirically, and therefore important to include in any taxonomy.

The term "fake news" was first used to describe satirical shows and publications (i.e., Daily Show, The Onion). For creators of such content, the concept meant made-up news, with the pursuit of entertaining others, and not for informing or deceiving. The first disagreement when defining fake news is if satirical publications should be included in the definition of fabricated news. Some scholars claim that satire should be left out of the new definition of "fake news" because it is "unlikely to be misconstrued as factual" and it is not created with the purpose of informing audiences (Allcott & Gentzkow, 2017, p. 214). However, others claim that it should be included because although it is legally protected speech, it could be misconstrued as telling the truth (Klein & Wueller, 2017). For example, in 2017, a satire site run by hoaxer Christopher Blair issued an apology for making their story "too real," after many were unable to detect its satirical nature (Funke, 2017).

The second disagreement when conceptualizing "fake news" is intentionality. Some scholars believe that for content to be considered fake, the content creator must have deceitful intent. For example, Allcott and Gentzkow (2017) and Conroy et al. (2015) argue that "fake news" should be defined as news articles that could mislead readers and are intentionally and verifiably false. This includes intentionally fabricated pieces and satire sites, but excludes unintentional reporting of mistakes, rumors, conspiracy theories, and reports that are misleading, but not necessarily false (Allcott & Gentzkow, 2017; Klein & Wueller, 2017). Such conceptualization leaves out mainstream media misreporting from scrutiny. As Mihailidis and Viotty (2017) explain, journalists face an information environment where the economic, technological, and sociopolitical pressures are combined with a need to report with speed, while engaging audiences in the process. This tension creates an environment where online news media become part of the problem of misinformation. This issue is best described in Benkler, Faris, and Roberts (2018). Findings of this extensive study illustrate how major news outlets such as *The New York Times, The Washington Post*, and *Associated Press* were involved in disseminating false information. For example, according to the authors, "Russian and right-wing political actors have been particularly effective at using document dumps, particularly hacked emails, to lure journalists into over-reporting" (Benkler et al., 2018, p. 358). Despite misreporting being unintentional, it is still an instance of untrue information disseminated via traditional as well as online media channels.

Finally, a third disagreement regarding "fake news" has to do with its conceptualization as a binary variable versus one that varies on a continuum. For example, the conceptualization of "fake news" as exclusively satire provides a binary differentiation between genres. It is either hard/mainstream news (based on real facts with the purpose of informing) or "fake news" (made-up stories with the purpose of entertaining). However, literature exploring "fake news" post-2016 election argues that "fake news" should be seen as a continuum because there is some online content that have biases, but do not fall directly into "fake news." As Potthast, Kiesel, Reinartz, Bevendorff, and Stein (2017) explain "hardly any piece of 'fake news' is entirely false, and hardly any piece of real news is flawless" (p. 4). Those who conceptualize "fake news" as a continuum include human fact-checking sites such as Snopes and Politifact that classify articles based on their degree of fakeness, rather than as absolutely true or false. For example, Politifact classifies "fake news" on a 6-point scale ranging from *true* to *pants on fire* and Snopes classifies in 12 categories including, true, false, mostly true, mostly false, mixture, misattribution, legend, and scam. Other human fact checkers include FactCheck.org, whose goal is to verify statements in transcripts and videos, and the Latin American Chequeado, that follows a similar classification system as Politifact.

It must also be noted that the vast majority of online fact checkers are based on corroboration with a database of verified facts. While they can be quite useful in identifying fabricated stories about established facts and also for training machine-learning algorithms, this method cannot help us determine the veracity of new, incoming information about developing stories, as is often the case with the recent crop of fabricated news surrounding elections, disasters, and mass shootings. Therefore, we need a more comprehensive view of fabricated news, one that not only checks on facts, but also linguistic characteristics of the story, its source(s), and the networks involved in its online dissemination. This is especially true given that when labeling and differentiating between problematic information, there is rarely a mutually exclusive characteristic because their meanings often have overlapping boundaries (Jack, 2017). The difficulties associated with human coding in such instances can be facilitated by the use of the probabilistic functioning of machine learning, where the machine is provided with input variables x to predict output y based on a series of classifications (Kotsiantis, 2007; Murphy, 2012).

Identifying fabricated information online before it becomes viral is imperative for maintaining an informed citizenry able to make decisions required in a healthy democracy. Using machine learning and including several types of indicators could provide a solution for the identification of fabricated information, despite the overlapping boundaries it has with other types of information. With this in mind, we propose an original taxonomy of online content, as a precursor to identifying signature features of fabricated news.

Taxonomy of Online Content for "Fake News" Detection

In our taxonomy, we identify eight categories of online content for the purpose of algorithm-based detection of "fake news:" real news, false news, polarized content, satire, misreporting, commentary, persuasive information, and citizen journalism. These categories are organized based on a combination of unique features derived and compiled from the various conceptual and operational definitions proposed for fabricated news through our meaning analysis. Such features include characteristics related to the message and its linguistic properties, its sources and intentions, structural components, and network characteristics. In the next section, we will first differentiate between real news and false news. Then, we identify online content that is not false news, but that could be misinterpreted by audience as false news. These types of online content are important to identify for the sake of building a taxonomy that has discriminant validity in ruling out content that is not false. Once identified, we can build algorithms to label the varied forms of news that exist between the binary categories of real and false, so that the reader can factor that in their consumption of such information or discourse. It will also serve to reduce reactance that is known to occur when readers are told that a piece of news which aligns well with their political beliefs is indeed false in a blanket manner. Providing a more nuanced labeling of partisan content, for example, without declaring it outright as false, can serve to balance the need for identifying content that is completely false and made-up and recognizing content in which truthfulness might be contested, as might be the case with partisan and persuasive content. This will also help enhance credibility of the algorithm and greater acceptance of its classification of different kinds of real and false news and the various shades in between the two.

What Is Real and What Is "Fake News?"

Real News. The first category of content in our taxonomy is of course real news. Although difficult to define, it can be understood through the journalistic practices that surround its creation. Some scholars argue against utilizing this approach because the notion of truth results from the subjective interpretation of reality such that news is an outcome of sensemaking by each epistemological community rather than an outcome of the newsroom (Giglietto, Iannelli, Rossi, & Valeriani, 2016; Waisbord, 2018). While we agree that professional journalism is not a perfect system, it served rather well in providing credible and objective information during the 20th century (Lazer et al., 2018). The social media ecosystem has changed the vertical orientation of truth governed by media outlets during these times to a horizontal relationship such that everyone can create and disseminate content. This has some advantages in the gatekeeping process, but it also comes with challenges (Waisbord, 2018). However, when it comes to false information transmitted online that can be potentially damaging (i.e., spread of false information in India leading to mob lynching of innocent people, health misinformation that can lead to damaging health outcomes, or made-up information for monetary gain or political destabilization), acknowledging the existence of truth is not only well-justified, but a critical need for a healthy democracy. Thus, disambiguating real and false information necessitates a grounding of truth. We understand this grounding as collective consensus, as proposed by Southwell et al. (2017).

Importantly as well, false news typically imitates real information in its form, but "not in organizational process or intent" as they "lack the news media's editorial norms and processes for ensuring the accuracy and credibility of information" (Lazer, et al., 2018, p. 1094). If the difference between real information and false information lies in such editorial decision making, then focusing on journalistic norms and rules to define real information is a reasonable approach to define truth for machine-detection purposes.

Real news includes hard news (breaking news) and soft news (less timely information; Shoemaker, 2017) and it is created through journalism defined as "the activity of gathering, assessing, creating, and presenting news and information" and abides by principles of verification, independence and obligation to report the truth (American Press Institute, 2017, para.1). The operational definition of this type of content is the pursuit of knowledge through verification based on reliability, truthfulness, and independence (Borden & Tew, 2007; Digital Resource Center, 2017). Another definition for real news is based on the news values that surround it (Shoemaker, 2017) and the

Table 1. Features of Real News.

Message and linguistic	Sources and intentions	Structural	Network
Factuality: • Fact-checked. • Impartial reporting. • Uses last names to cite. Evidence: • Statistical data, research-based. Message quality: • Journalistic style. • Edited and proof read. Lexical and syntactic: • Frequent use of "today." • Past tense. Topical Interest: • Conflict. • Human interest. • Prominence.	Sources of the content: • Verified sources. • Quotes and/or attributions. • Heterogeneity of sources. Pedigree: • Originated from a well-known site/ organization. • Written by actual news staff. Independence: • Organization associated with the journalist.	URL: • Reputable ending. • URL has normal registration. About Us section: • Will have a clear About Us section. • Authors and editors can be verified. Contact Us section • E-mails are from the professional organization.	Metadata: • Metadata indicators of authenticity.

techniques to decide newsworthiness (Trilling, Tolochko, & Burscher, 2017). These characteristics include proximity and scope, timeliness, social and personal importance, topic-specific characteristics, and objectivity. For the purpose of identification of real and false information, only the latter two are relevant operational definitions given that the former can vary based on geographical regions and other personal factors. In essence, these are not characteristics of real news, but rather indicators of where and to whom a news story may have higher popularity or importance.

These definitions can further be deconstructed into specific features that differentiate real news from other categories of information (see Table 1). First, the objectivity of a message can be assessed through its factuality and impartiality achieved through fact-checking and quote verification (Borden & Tew, 2007; Shoemaker, 2017). We acknowledge that factuality is a highly debated topic. While some agree it is needed to maintain democracy, others argue that it can create coverage that supports elites (Shoemaker, 2017). There are several online services available for fact-checking and quote verification. For example, Storyzy (2017) is an online quote verifier able to decipher if a quote is authentic. As many as 50,000 new quotes are entered daily, making it a viable resource for quote verification. Importantly, quote verification is only one part of fact-checking. An overall assessment of the claims in an article is essential. To this effect, a knowledge-based paradigm of machine detection through information retrieval, semantic web, and linked open data can be employed (Potthast et al., 2017). For example, an analysis can be done to "extract and index factual knowledge from the web and use the same technology to extract factual statements from a given text in question" (Potthast et al., 2017, p. 2).

Although factuality is an important component to assess objectivity, impartiality should also be assessed. Impartiality can be understood as the use of tools to achieve objectivity and include the inclusion of sources, attributions, and a balanced coverage (Shoemaker, 2017). Real news uses credible sources and attributes them throughout a news story. Source characteristics can help identify the reliability and veracity of real news. For example, real news typically reports directly from official sources and tries to present a list of heterogeneous sources with different sides of an issue or event (Shoemaker, 2017).

Other message features of real news include stylistic indicators such as the adherence to journalistic style of writing (Frank, 2015), the use of an inverted pyramid style of writing, the absence of storytelling characteristics (Shoemaker, 2017), as well as lexical and syntactical structures (Argamon-Engelson, Koppel, & Avneri, 1998) that are characteristic of real news content. Message characteristics can also be topical in nature. According to Shoemaker (2017) and Trilling et al. (2017), news is often controversial in nature and includes conflict and human interest. Finally, real news has structural characteristics which include its independence, or lack of affiliation with an interest group, as well as accountability (American Press Institute, 2017).

Although the characteristics presented for real news represent the ideal of news, the 24-hour cycle and the social media environment might make these unrealistic characteristics. As Shoemaker (2017) explains, journalists nowadays are not as strict to follow standards and news values. However, in the social media environment where everyone can create and share information, good journalism is more important than ever not only to maintain an informed citizenry but also to increase trust in media.

It is also important to acknowledge that the purpose of journalism is to provide citizens with information necessary to decide on the needs of the state. To this purpose, journalists select what information is newsworthy and frame it accordingly. As McCombs (2005) explains, there are several players in the process of agenda setting. This includes journalists, editors, and other external organizations. Although the process of news development may be seen as the creation of reality that might be somewhat subjective, this does not mean that the content created is made-up as in false news (Tandoc et al., 2018).

An example of real news is a *BBC* news article about Venezuela's crisis.[1] This story has several features of real news. The corresponding statistics can be fact-checked and is written in journalistic style (message characteristics); it uses verified sources and it is written by a well-known media organization (characteristic of the source and intentions); and the URL has a reputable ending and a clear "about us" section (structural features).

False News/Hoaxes. False news, on the other hand, is defined as information that is intentionally false, and are often malicious stories propagating conspiracy theories. Although this type of content shares characteristics with polarized and sensationalist content (described later in this article), where information can be characterized as

Table 2. Features of Fabricated News.

Message and linguistic	Sources and intentions	Structural	Network
Factuality: • Not factual. • One-sided reporting. Message quality: • Grammar, spelling or punctuation mistakes. • Does not adhere to journalistic style. • Cites first names. Lexical and syntactical: • Present tense verbs. Rhetorical elements: • Discrepancies or omissions. • ”Spectacle” and narrative writing. • Emotionally charged. • Use of hyperboles. • Common man appeals. • Arguments from authority. • Ad-hominem attacks. • Demonizing the out-group. • Conspirational reasoning. • Logic flaws. Headline • All CAPS and exclamations. • Misleading and clickbait headlines. Sound bites: • Editing soundbites to create sensationalism. Photos/videos: • Altered pixel structure. • Shadows, reflections, and perspective distortions. • Use of photos out of context.	Sources of the message: • Unverified sources. • No quotes or made-up quotes. • No source attribution. Intentionality: • Intentionally false. • Revenue purpose. Independence: • Source of origin is not reputable. Pedigree: • Originated in an obscure site or social media post. • Not vetted by mainstream media.	URL: • Not reputable ending (.com.co). • Recently registered URL. • Designed to look like established site. • Ephemeral site. About Us section: • Does not have information about the editor or listed owner. Contact Us section: • E-mail is a “personal” address Uncommon journalistic practices: • Provide a Free PDF version. • Asks users to send their stories for publication. Comments: • Asks users to comment to access an article. • Red flag if many users say it is false.	Personalization and Customization: • Circulated via social media where content is tailored to user interests. Social media shares: • Often shared through social media by mutual friends or preidentified accounts. Author: • Written by bots and algorithms. Metadata: • Metadata indicators of authenticity.

highly emotional and highly partisan (Allcott & Gentzkow, 2017; Howard, Kollanyi, Bradshaw, & Neudert, 2017; Potthast et al., 2017), it differs in important features (see Table 2).

First and foremost, false news stories are not factual and have no basis in reality, and thus, are unable to be verified (Allcott & Gentzkow, 2017; Cohen, 2017). Furthermore, false news differs from polarized content in structural components. For example, false news often originates from ephemeral sites created for ad revenue purposes. In fact, many of the fabricated sites of the 2016 elections were later identified as having fairly recent registration domains and foreign locations (Silverman, 2016; Soares, 2017). For example, the site *Ending the Fed*, responsible for disseminating four of the top 10 false election stories identified by Silverman (2016), had a domain registered in March, 2016, a mere 8 months prior to the elections. A well-established news organization will not have created their site a couple of months ago. Nevertheless, the aforementioned stories by *Ending the Fed* generated nearly 180,000 more engagements on Facebook than the top four election stories from *The Washington Post* (Silverman, 2016). As Dempsey (2017) explains, fabricated sites can make a lot of revenue through online ads by simply driving users to their sites so "they may not care what the content even says or how misleading the headline is, as long as it is attracting eyeballs to their paid-per-click pages." Because the main purpose of this type of content is ad revenue, message and source characteristics are often neglected. It is not uncommon for a false news article to have unverified quotes, emotionally charged linguistic markers, spelling and grammar mistakes, and inaccurate pictures. A quick reverse-Google search, for instance, can reveal if a displayed picture occurred prior to the event it claims to be reporting about (Davis, 2016).

Finally, network features are especially salient for fabricated sites. Because these websites intentionally publish deceptive and incorrect information for financial gain, they rely on social media to engage audiences. As such, false stories are often circulated via social media without being reported by mainstream/professional news organizations (unless they cover this story for the express purpose of alerting users about their false nature). Benkler et al. (2018) call this networked propaganda, referring to the social media architecture that lends itself to mass dissemination of false information. Thanks to algorithms and recommendation systems, users often receive content based on their own prior exposure and those of like-minded others. This can narrow the universe of stories read by any one individual and lead to the creation of "echo chambers" (Sunstein, 2002) and "filter bubbles" (Pariser, 2011).

As Howard et al. (2017) elucidate, "both fake news websites and political bots are crucial tools in digital propaganda attacks—they aim to influence conversations, demobilize opposition and generate false support" (p. 1). For example, a recent study revealed that social bots in Twitter served to amplify the dissemination of content coming from low-credibility sources, content such as conspiracy theories, false news, and junk science, suggesting that "curbing social bots may be an effective strategy for mitigating the spread of low credibility content" (Shao et al., 2018, p. 5-6). One of the strategies suggested to identify social bots includes the design and development of machine-learning algorithms.

Research suggests that social bots specifically aid dissemination at early stages of virality and target people who might be vulnerable to this information and are thus likely to reshare the content (Shao et al., 2018). Following this premise, Potthast et al.

(2017) recommend following a context-based paradigm of deception, through social network analysis assessing the spread of a particular piece of information. Hamilton 68, a site that tracks Russian propaganda and disinformation on Twitter, is an example of network features in action. As the site describes, the dashboard monitors the activity of accounts assembled based on a 3-year analysis tracking disinformation campaigns and identifying both humans who shared such content and the use of bots to "boost the signal of other accounts" (Berger, 2017, para. 10). Another recommendation proposed by Conroy et al. (2015) is the use of network approaches to detect deception by using elements such as message metadata to provide a cumulative measure of detection.

An example of false news is a story reported during Hurricane Harvey about looters targeting Trump supporters.[2] This story is not fact-checked (and has been debunked by factcheckers[3]), it is emotionally charged (message characteristics); the tag line of this outlet: "airing out America's dirty laundry" reveals the source is not a mainstream news organization; and the sources within the story include Twitter posts, many later identified as belonging to accounts with a history of disseminating misinformation and deleted since (structural and sources characteristics).

Myths About False News: The Gray Area

Even though some content posted online may indeed be misleading, it is important to not confuse them with false news or real news. The goal of developing algorithms for fabricated news detection is not to impinge on users' right to express their opinions, or in journalists' endeavors, but to stop dissemination of false information. However, in order to achieve this goal, we need to identify the types of online content that are often misattributed to be false news or real news.

Commentary, Opinion, and Feature Writing. The first type of content to be aware of is commentary and feature writing. Although these are typically pieces written by mainstream and traditional outlets, many often confuse it with hard news. Commentary and other similar editorial pieces are different from real news in that the journalist does not abide by principles of opinion-free reporting typically seen in hard news stories (Digital Resource Center, 2017; Padgett, 2017). Yet opinion journalists are well within their rights to express opinions. Their job is to select facts to form an argument and adhere to the Society of Professional Journalism code of conduct when doing so (Digital Resource Center, 2017). Nonetheless, the 24/7 news cycle exacerbates the need to fill time, and thus commentators are more often featured, blurring the lines between news and commentary (Turner, 2017). Importantly, commentary should not be confused with an assertion because its emphasis is on providing conclusions based on evidence, whereas an assertion (typically seen in polarized and sensationalist content) occurs when something is declared without the necessary evidentiary basis (Digital Resource Center, 2017).

Identifying commentary, thus, requires differentiating between both real news and assertions (Howard et al., 2017; see Table 3). This can be done based on opinion journalists' adherence to the code of professional conduct of the Society of Professional

Table 3. Features of Commentary.

Message and linguistic	Sources and intentions	Structural	Network
Factuality: • Based on facts and evidence. • No misrepresentations. Evidence: • Statistical data, research-based. Rhetorical elements: • Emotionally charged. • Narrative writing. Message quality: • Journalistic style. • Edited and proof read. Lexical and syntactic: • Frequent use of "should."	Sources of the content: • Written by actual news source. Pedigree: • Originated from a well-known site/ organization. Independence: • Organization associated with the journalist.	URL: • Reputable ending. • URL has normal registration. About Us section: • Will have a clear About Us section. • Authors and editors can be verified. Contact Us section: • E-mails are from the professional organization. Labeling: • Labeled as commentary, editorial, analysis.	Metadata: • Metadata indicators of authenticity.

Journalists and based on the Verification, Independence, and Accountability principles of journalism. Following the SPJ and VIA principles, specific characteristics of message and structure can be identified to detect commentary. Features of structure include, at the outset, the labeling of commentary as such. For example, in professional news outlets, opinion pieces are located in the editorial, opinion section, or other such markers of the category of news that one is consuming. Additionally, commentators should remain free from activities than can damage their reputation and should remain independent. The latter can be assessed through structural features allowing the reader to assess the partisan associations, if any, of the commentator (Digital Resource Center, 2017). On the other hand, features of the message or linguistic structure include the presentation of opinions based on facts, evidence, and expert opinions (Borden & Tew, 2007). Furthermore, the message in commentary tends to be more emotionally charged and opinion-based compared with real news—features that can be identified using linguistic markers of the message (Argamon-Engelson et al., 1998; Borden & Tew, 2007).

An example of a commentary is the article written in *The New York Times* about Ilhan Omar.[4] The article uses the word "should" throughout the article and includes other opinion linguistic markers (message characteristics), it is written by a well-known organization (characteristic of the source and intentions), but is clearly labeled as an opinion piece (structural characteristic).

Misreporting. Related to real news and commentary is misinformation, or unintentional false reporting from professional news media organizations. Even though the intention of professional journalists is not to be deceitful, misreporting can sometimes

Table 4. Features of Misreporting.

Message and linguistic	Sources and intentions	Structural	Network
Factuality: • Not-fact checked. • Uses last names to cite. Message quality: • Journalistic style. • Edited and proof read. Lexical and syntactic: • Past tense.	Sources of the content: • Unverified sources. Pedigree: • Originated from a well-known site/organization. • Written by actual news staff. Independence: • Organization associated with the journalist.	URL: • Reputable ending. • URL has normal registration. About Us section: • Will have a clear About Us section. • Authors and editors can be verified. Contact Us section: • E-mails are from the professional organization. Labeling: • Presence of a correction or retraction when identified.	Metadata: • Metadata indicators of authenticity.

occur. Thus, it is important for it to be included in a taxonomy of online content for the purpose of false news detection. That being said, this content should not be confused with blatantly false reports created with the intention of being deceitful. Misreporting is an example of misinformation "defined as false, mistaken, or misleading information" and not disinformation "the distribution, assertion, or dissemination of false, mistaken, or misleading information in an intentional, deliberate, or purposeful effort to mislead, deceive, or confuse" (Fetzer, 2004, p. 231).

Misreporting can be distinguished from real and completely false news (see Table 4). The first important set of features is that of sources and intentions evaluated based on reliability and objectivity. For instance, misreported articles will typically not report straight from sources and some of its quotes might not be verified. Similarly, this content can be evaluated based on message and linguistic features such as its one-sided reporting, and factual inaccuracies.

Finally, structural elements demarcated based on principles of accountability and independence can also be identified to differentiate this content from disinformation. For instance, although the article contains one-sided or inaccurate reporting, the journalist shows accountability by providing his or her name and affiliation with a news organization. Similarly, the URL of the site provides information about the news organization and its independence from external groups that might benefit from the misinformation in question. And, when it is pointed out, the organization will issue a retraction or correction.

An example of misreporting is a story published by *The New York Times* about the new tax code in 2018.[5] The article reported on a hypothetical couple whose owed taxes would increase by approximately 3,000 dollars, when in reality it would have decreased by 43 dollars. The story is from a reputable organization and has a reliable URL (structural and source characteristics), but when fact-checked it was shown to have errors, and the organization issued a correction (message characteristics).

Polarizing and Sensationalist Content. The next type of online information is polarized and sensationalist content. This content is not completely false but is characterized by its "goodness-of-fit with a particular ideology" (Berghel, 2017, p. 3). Importantly, although objectivity is not the goal of polarized or partisan media, authors might still assert truth through arguments that justify their own position (Shoemaker, 2017). According to Pennycook and Rand (2017), individuals fall for misinformation because they fail to think analytically when faced with misinformation. This is particularly true with information that agrees with their prior knowledge and beliefs (Bode & Vraga, 2015; Rich & Zaragoza, 2016). Both message features and source/intent features can help recognize when content is aligned with a particular ideology. For example, in Potthast et al. (2017), researchers were able to differentiate between hyperpartisan, mainstream, and satire based on its style of writing through unmasking style-based categorization using computational techniques. However, they were unable to differentiate between left- and right-wing style of writing, suggesting that at least stylistically both have a lot in common. Findings of this study demonstrate that features related to the style of content can be successfully employed to distinguish between mainstream news and polarizing news.

A common strategy utilized by polarized content is the use of highly emotional and inflammatory content. These assertions typically lack evidence and are based on appeals to emotion and preexisting attitudes (Allcott & Gentzkow, 2017; Digital Resource Center, 2017). Operational definitions for this type of content include the presentation of divisive, inflammatory, and misleading information to manipulate the audience's understanding of an event or issue (Howard et al., 2017). This is problematic because when feedback about an event or object is limited, users tend to judge quality of content based on previous experiences and prefer confirmatory reporting as it provides psychological utility (Allcott & Gentzkow, 2017). Furthermore, assertions typically are constructed based on implied information, or self-generated deduction, which is stronger because it activates previous schemas and scripts (Rich & Zaragoza, 2016). Another reason why people are attracted to sensationalist content is the events being covered imply danger (Ng & Zhao, 2018). Ng and Zhao (2018) explain that sensational news stories "satisfy the human survival instinct to detect environment threats" (p. 4). As such, people read and click more on headlines that contain alarm words, such as "accident" or "death"—a message feature that serve as an alarm system to detect possible threats and respond to them accordingly. Furthermore, people read and click more on alarm words when these are paired with prosocial words such as "empathy" and "help" (see Table 5).

Other message features to help identify sensationalist and polarized content include the use of ad-hominem attacks, hyperboles, and other attention-grabbing techniques; as well as its lack of evidence to support claims and display of inaccurate pictures and soundbites (Borden & Tew, 2007; Davis, 2016; Howard et al., 2017; Najmabadi, 2016; Reilly, 2012; Roozenbeek & van der Linden, 2019). Stylistically, the headlines of

Table 5. Features of Polarized and Sensationalist Content.

Message and linguistic	Sources and intentions	Structural	Network
Factuality: • Not completely factual. • One-sided reporting. • Degree of fit with a political agenda. Message quality: • Lacks evidence. • Excessive capitalizations. • Use of alarm words. • Use of prosocial words. Rhetorical elements: • Generalizations. • "Spectacle" and narrative writing. • Emotionally charged. • Ad-hominem attacks. • Logic flaws. Headline • All CAPS and exclamations. • Misleading and clickbait headlines. Sound bites: • Editing soundbites to create sensationalism. Visuals: • Extreme use of static and moving visuals.	Sources of the message: • One-side sources. Independence: • Source of origin is not reputable. Pedigree: • Polarized source of origin.	About Us section: • Speaks ill of media. • Describes itself as leaning toward a political side.	Personalization and Customization: • Circulated via social media where content is tailored to user interests. Social Media Shares: • Often shared through social media by mutual friends or preidentified accounts. Metadata: • Metadata indicators of authenticity.

polarized content are clickbait style, accentuating negative content and exaggerating the article. There are already computational models under development for detection of clickbait. For instance, Shu, Wang, Le, Lee, and Liu (2018) propose a Stylized Headline Generation system that can generate headlines for training, while improving classification for supervised machine learning. It is important to note, however, that the use of clickbait headline by sensationalist content does not make the headlines necessarily false. As Ecker, Lewandowsky, Chang, and Pillai (2014) explain, clickbait content can be misdirected such that the information is technically true, but "misleading in substance" (p. 324).

Another strategy to determine goodness of fit are features of the sources/intent. For example, sensationalist content typically includes only one-sided sources that are often intended to provoke (Potthast et al., 2017; Wardle, 2017). Structural features can also be giveaways of their polarized nature. For example, often in the "about us" section, these sites will speak ill of mainstream media. Finally, network features can shed light on whether the content was personalized to be received by certain individuals or networks of individuals with particular partisan affiliations or views. User tailoring, for instance, facilitates the creation of echo-chambers, which many argue reduces exposure to counterattitudinal information and therefore the likelihood of perceiving content as partisan (Dylko et al., 2017; Sunstein, 2002). As such, features of the network are essential when identifying polarizing content (see Table 5).

An example of polarized content is an article[6] written by the site deepleftfield.info about Trump's alleged criminal charges. The article has a mix of facts and opinions (message characteristic), the source is a polarized source of origin as classified by fact checkers such as mediabiasfactcheck.com (characteristics of sources and intentions), and has a clear political ideology as expressed in their byline "the place for progressive political debate" and URL (structural characteristic).

Citizen Journalism. In the digital world, reporting has become a collaborative and participatory practice. As Downie and Schudson (2009) explain, news gathering is not only conducted by the traditional newsroom, but also by freelancers and citizens at large. Although these nontraditional sources can provide important information about breaking news and be eye-witnesses to events otherwise not covered, some of their reports can be unverified and inaccurate (Conroy et al., 2015), or simply personal commentary for sharing opinions, photos, observations, and activities (Downie & Schudson, 2009). There are two different categories of citizen journalism. The first includes blogs and websites from citizens, civil societies, and organizations with content originally created by such users (Howard et al., 2017). Features to distinguish this content from other online information include message aspects such as not adhering to journalistic style of reporting and verification. Additionally, its contents are more emotionally driven and subjectively reported. Last, there are essential structural components. For instance, the URL and "about us" section are giveaways of the site being a blog, personal site, or a site specifically meant for citizen reporting.

The second category of citizen journalism refers to subsites of professional journalism sites providing a forum for citizen reporting (e.g., *CNN's* iReport). These sites are often focused on first-hand eye-witness accounts of breaking news. Message features to identify this type of content is similar to that of personal sites and blogs. For example, the content is typically a video recording, and when written, it lacks journalistic style of reporting. This content can also be identified through structural features. Even though this content is technically a subsection of a professional media site, this subsection is identified as content developed by users (see Table 6).

Table 6. Features of Citizen Journalism.

Message and linguistic	Sources and intentions	Structural
Message quality: • Does not adhere to journalistic style. Lexical and syntactic: • Past tense. Modality: • Often in video format.	Sources of the content: • Unverified sources. Pedigree: • Originated from audience member. Independence: • May reflect creator's political and organizational affiliations.	URL: • Site is a blog or personal site. • Subsite of news organization labeled as audience created. Contact us section • Personal e-mails from bloggers.

Although the clearest example of citizen journalism is *CNN*'s iReport, there are several occasions worldwide where citizens have documented news events. For example, in 2004, blogs and social network groups provided information about the tsunami that hit Southeast Asia (Adweek, 2009). Posts were typically in video format (message characteristics), originated from audience members (characteristics of sources and intentions) and the online sites were blogs or social media posts (structural characteristics).

Satire. Another source of content that can be found online is satirical news, defined as an intentionally false story meant to be perceived as unrealistic, that uses journalistic style as a parody of the style, to mock issues and individuals in the news, or to disseminate a prank (Frank, 2015; Reilly, 2013). As Balmas (2012) explains, satire derives from hard news and not from reality; and it is intended to be perceived as unreal. Even though some scholars argue that it should not be included in the description of "fake news" as its contents are expected to be humorous (Allcott & Gentzkow, 2017; Borden & Tew, 2007), it is important to categorize it explicitly so that automated detection of false news can identify it for what it is, avoiding misclassifying it as either real or false. This is especially important because news consumers are not always aware of its satirical nature, especially when they are forwarded such "news" via informal social media and messaging channels.

There are three basic set of characteristics of satirical content (see Table 7). First, it needs the understanding of what it means to be an authentic and legitimate news practice (Baym, 2005). In other words, although satirical news may adhere to journalistic style of reporting, it does so in a humorous way rather than abiding by principles of truthfulness and verification required for journalistic practices. This satirical style of writing can be assessed through message and linguistic features such as its adherence to journalistic style of writing. However, it differs from real news in that it is not fact-checked (message feature); it often has amateur mistakes such as spelling mistakes, grammar mistakes, and the overuse of vague expressions such as "more and more" (message feature); and it either includes fake sources, or made-up quotes (feature of the source and intention). Previous research suggests that stylistic features can be employed to differentiate satire from false news and real news. For example, Potthast

Table 7. Features of Satire.

Message and linguistic	Sources and intentions	Structural
Factuality: • Not-fact checked. Message quality: • Journalistic style of writing but may have amateur mistakes. • Overuse of typical journalistic expressions. • Grammar, spelling, and punctuation errors. Rhetorical elements: • Use of hyperboles. • Exaggerations. Headline: • Clickbait headline. Context: • Humorous content.	Sources of the content: • Fake sources. Pedigree: • Originated from a satire site.	About Us section: • Label as satire. • Self proclamation. Publication date: • Postdated. • Dated April 1.

et al. (2017) used unmasking categorization techniques based on stylistic features, including dictionary features of frequency of words, readability scores, ratio of quoted words, among others, to discover that “the style of fake news has much more in common with that of real news than either of the two has with satire” meaning that “it should be fairly easy to discriminate between the two [satire and real news]” (p. 8).

The second characteristic of satire is the type of website that is created for this purpose. Frank (2015) identifies five types of satirical websites that can be identified based on message quality, amateur mistakes, and clickbait headlines illustrated by excessive capitalizations and use of hyperboles (message features). Finally, a third feature of satire is that it uses real news as its reference point (Balmas, 2012). Structurally, satirical content is labeled as such in the “about us” page and is often satirical in itself as illustrated by self-proclamation. Additionally, the date of publication of these articles can have inconsistencies (Davis, 2016; Ojala, 2017; Reilly, 2013). For example, a well-known prank by the “yes men” group published a special edition of *The New York Times* depicting a utopian United States (Reilly, 2013). Those who believed it did not realize it was postdated, among many other indicators of satire.

The most popular examples of satire are the articles written by the Onion. For example, the article about the 2019 U.S. government shutdown[7] is written in journalistic style, but with a humorous spin (message characteristics); the lead source quoted in the article is made up, as can be determined by a Google search; and the “about us” section includes self-proclamation and is a satire in itself.

Persuasive Information. Another type of content to consider is persuasive information, which can be further decomposed as native advertisement and promotional content.

Table 8. Features of Persuasive Information.

Message and linguistic	Sources and intentions	Structural
Message quality • Content directly associated with brand. • PR for candidate or agenda. Rhetorical elements: • Narrative writing. Factuality: • One-sided arguments. • Balanced testimony of untruth values. Evidence: • Flawed statistical data. • No replication. Lexical structures: • Second person statements.	Sources of the content: • If it has sources, it is associated with brand or groups related to the agenda. • Sources from science dissidents. • Financial-gain intention. Independence: • Author is an organization or promotional entity.	Labeling: • Labels indicating paid promotion in social media. URL: • Site is from political, governmental, or another public agency. About Us section: • States goals of the site.

Although conceptually both subsections are distinct, they share their persuasive intent (Jack, 2017). Thus, we combine them into one category for the sake of parsimony.

The first type of persuasive information is native advertisement, defined as promotional material of persuasive intent, often masked as a news article (Frank, 2015). Because this type of content takes the form of the platform it is distributed on, it is not surprising that users might treat it as true, even when it is clearly labeled as promotional. In a Stanford study, students were unable to discern between sponsored posts, real news, and biased news. According to the report, over 80% of students believed the native advertisement was real news, despite it being labeled as sponsored (Stanford History Education Group, 2016).

Specific message features to help answer these questions include the presence of a logo from the company indicating authorship or sponsorship and verification of content through cross-checking (see Table 8). Furthermore, features of the source include identifying the author of the content as well as the sources within the article and their relationship with the author or company of interest. For instance, even if it is written as a narrative, the native ad will promote their product throughout the story being told (Carlson, 2014). Finally, structural elements include the labeling of the content as promotional, sponsored, or paid content. For example, a native advertisement in Facebook can be identified by a "sponsored" tag. Similarly, paid content in Instagram will be tagged as "paid partnership" or will include hashtags like #sponsored or #ad.

Yet another type of persuasive information is promotional content, both political and nonpolitical. Professional political content includes information from government, political parties, or public agencies, as well as information from experts (Downie & Schudson, 2009; Howard et al., 2017). It is important to include this type of content because it reflects the agendas of particular political parties or organizations. Although not false news per se, they do not abide by objective two-sided reporting, and thus should not be confused with real news. Promotional content can take the form of official content produced by a political party, candidate's campaign, government, or public agency; and white papers or policy papers from think tanks.

Official content can be identified though message features advancing the agenda of a candidate; for example, statements like "we will" or "the political party will" are typical. Furthermore, the content will typically lack sources, or feature only sources related to the specific group or agenda (see sources and intentions features in Table 8) and will typically be published on a URL related to the organization (see structural feature in Table 8). Think tank papers are similar to political party content. Features of the message would differ in that it will not necessarily advance an agenda but might possess opinion linguistic markers (see message features in Table 8) and will often cite academic and research sources rather than specific individuals (see sources and intentions feature in Table 8).

In the promotional content category, it is also important to recognize content outside of the political realm such as that emerging from organizations that promote an idea or specific agenda. For example, powerful interests can push their agenda by attempting to persuade readers in a certain direction, such as the efficacy or lack thereof of alternative medicine. As Sheble (2018) explains, when such findings align with powerful external interests (e.g., pharmaceutical companies) they "may influence public perception and cast doubt on research findings even when there is broad consensus within science" (p. 157). Another type of promotional content is public-relations materials that tend to be published as news (Tandoc et al., 2018). Although press releases are typically e-mailed to media contacts in hopes that they will incorporate it in their news report, it has become increasingly popular for newswire services to publish them directly online so that they can be found by media organizations and the general public alike (iReach, 2018; Tandoc et al., 2018).

Message characteristics that can be used to identify promotional content are its one-sided arguments produced by third party authors (Tandoc et al., 2018), and the flawed analysis of statistical data and replication (Crocco, Halvorsen, Jacobsen, & Segall, 2017; McElreath & Smaldino, 2015; see Table 8). Nevertheless, McElreath and Smaldino (2015) elucidate the importance of understanding false positive rates when assessing reliability as well as the controversy of publication bias in the research community. Another message characteristic of promotional content is the use of the second person perspective. According to Cook (2001), the use of the word "you" is "one of the most distinctive features of advertising" and its use is the "most divergent from the uses of other genres" (p. 157). Finally, sources/intent can be analyzed by the inclusion of arguments from known dissidents when the content promotes a particular side or issue in an agenda (Jones, 2002; Sheble, 2018), as well as its self-promotion for financial gain in the case of news releases.

An example of this type of content is a native advertisement in *The Guardian* providing steps to end homelessness.[8] The article is written in narrative writing and is associated with the sponsor "stand together against poverty" (message characteristics), the sources within the article are associated with the brand (characteristic of sources and intentions), and the article is tagged as "advertiser content" in the site's upper left corner.

Features or Indicators of "Fake News"

Throughout this explication, initial features or indicators of "fake news" were identified based on a meaning analysis of descriptions found in academic articles, trade publications, newspaper, and magazines, among others. These features are decomposed as features of the message or linguistics, features of sources and intentions, structural features, and network features and can be useful when assessing online content based on the proposed taxonomy.

For example, through our taxonomy, we have identified an array of message features. Message quality elements such as punctuation errors or spelling mistakes are alerts of the possibility of an article to be false. Similarly, lexical components such as the word "should" or "you" can indicate that the content is possibly an opinion piece or promotional content. Visual components are also included, as a difference in pixel structure or a reverse image search can assess the realness or fakeness of an image.

Features of sources and intentions include sources within the message as well as the source of creation of the content. For instance, if an article does not attribute its sources, or these are not verified, it is a red flag or indicator that the content might be false. Likewise, features of the source include the site of origin and pedigree. An article coming from an obscure site or social media post is more likely to be false.

Structural features can also provide means of identifying type of content. Typical structural features of fabricated sites is the URL, often mimicking that of a traditional outlet, but ending in .com.co (e.g., www.abc.com.co), as well as a fairly recent date of registration. Other indicators include the "contact us" section, which provides a personal e-mail account rather than one of a reputable company.

Finally, features of the network are related to the dissemination of articles and structure of the technology that allows for this to occur. For example, false news is often shared by pre-identified accounts and our network of family and friends. Because companies like Facebook and Google keep their algorithms a trade secret, features of the network are more difficult to be identified, yet such features are important for better understanding dissemination of fabricated content.

The different features are meant to be indicators or red flags of fakeness, and each piece of online content should be analyzed in terms of all the features to determine where it belongs in our taxonomy. Importantly, these indicators or red flags are not, in and of themselves, deterministic of a type of content. Instead, the combination of several indicators can enable an algorithm to estimate the relative probability of a piece of content belonging to one of the eight categories in our taxonomy. For example, real news sometimes also has punctuation or spelling mistakes, nevertheless false information is more likely to possess these characteristics. Likewise, well-known credible news sources sometimes get facts wrong, but the likelihood of false news is higher if the source is an unknown URL. Because there are many overlapping factors, we identified different types of characteristics to assess the content in a holistic manner and calculate the likelihood of content belonging to a particular category in our taxonomy.

For example, the different features in our taxonomy can be built into sequences of yes/no questions (see e.g., in Table 9) conditioned on input features to build

Table 9. Example of Indicators of "Fake News" Converted Into Yes/No Questions for Decision Tree Algorithm.

Types of content	Fact-checked	Emotionally charged	Source verification	Registration inconsistency	Site pedigree	Narrative writing	Humor
Real news	Yes	No	Yes	No	Yes	No	No
False news	No	Yes	No	Yes	No	Yes	No
Polarized content	No[a]	Yes	No	No	No	Yes	No
Satire	No	No[a]	No	No[a]	No	No	Yes
Misreporting	No	No	No	No	Yes	No	No
Commentary	Yes	Yes	Yes	No	Yes	Yes	No
Persuasive information	No	No[a]	No	No[a]	No[a]	Yes	No
Citizen journalism	No	No[a]	No	No[a]	No[a]	No[a]	No

[a]Feature not available, tagged as "No" for the purpose of this exercise.

various supervised learning models such as Naïve Bayes, Linear Regression, or Support Vector Machines (Cortes & Vapnik, 1995; Maron, 1961). Among many popular supervised learning models, the Decision Tree model (see Figure 1) is one of the simplest learning models, wherein it is intuitive to interpret the classification result (Quinlan, 1986). The combination of features will indicate where in the taxonomy a particular piece of content will best belong, based on probabilistic calculations. For instance, as illustrated in Table 9, if a piece of content is not fact-checked, emotionally charged, written in narrative style, sources not verified, has inconsistencies with the registration date, and comes from an unknown source, then it is more likely to be flagged as "false news" than if the content is emotionally charged and written in narrative style, but is fact-checked, uses verified sources, comes from a known source, and does not have registration inaccuracies (in which case, it would be categorized as "commentary").

Importantly as well, Figure 1 illustrates a Decision Tree model using the subset of features listed in Table 9, assuming that the task is to distinguish eight different articles belonging to eight different types. Despite overlapping features, the system is able to classify the eight types of content by picking specific features that help identification in each instance, to reach a conclusion. Furthermore, as is common with machine learning output, performance accuracy is calculated. This helps further assess the certainty with which we can classify a piece of content in a particular category.

This Decision Tree implementation is an example of how the proposed taxonomy can be used in machine learning detection using a subset of characteristics identified in this explication. The use of machine learning for false news detection is still at an early stage. Future study of fabricated news characteristics can expand on the list of features we have proposed and test them, as we search for distinguishing features for including in such an algorithm. Through this testing, researchers can assess the

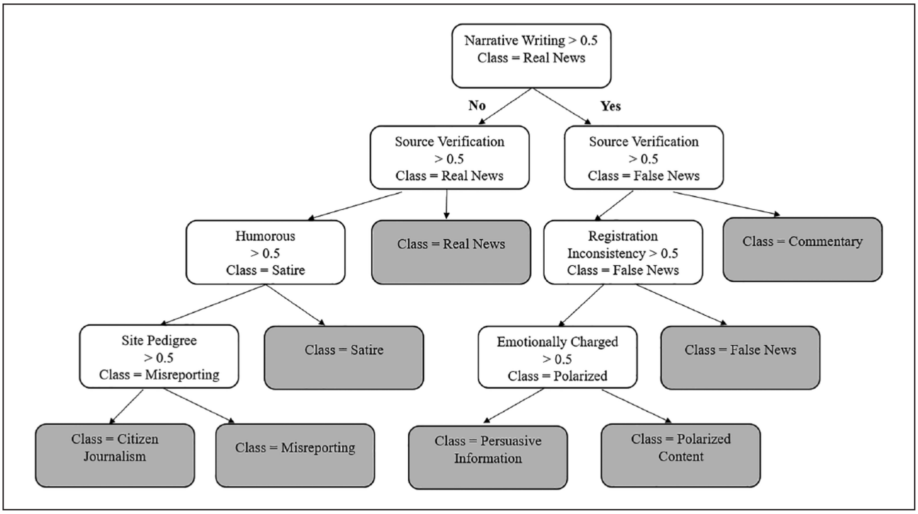

Figure 1. Decision Tree example using the subset of characteristics listed in Table 9.
Note. Decision Tree works like a flow-chart in which each decision node (represented by the white squares) is a binary "test" on an attribute or feature of false news (i.e., narrative writing, source verification); each branch or connecting line represents the outcome of the test (Yes/No); and each leaf of the tree, or gray box, reveals the class label (or the decision taken after computing the attributes). In this example, the initial node tests if the probability of an article being written in narrative style of writing is more than 0.5. There are two possible branches or outcomes. The "yes" branch indicates that the probability that the article is written in narrative style of writing is indeed more than 0.5. The "no" branch, on the other hand, indicates that the probability was lower than or equal to 0.5. Assuming that the outcome is "no," the next node or "test" would be that of source verification. Again, there are two possible outcomes or branches. If the probability of the article of using verified sources is less than or equal to 0.5, the outcome of the test will be "no." On the other hand, if the probability that the sources of this article are verified is higher than 0.5, the branch or outcome will be "yes," leading the system to its first decision, or leaf, classifying the article as real news. However, if the probability is less than or equal to 0.5, the outcome will be "no," in which case the content will then be assessed for how humorous it is. Following the same logic, if the article has a higher than 0.5 probability of being humorous, the outcome will be "yes," and the system will classify the content as satire. Importantly, this exercise is not deterministic in nature. Various Decision Tree algorithms can be used and tested to help us create similar decision trees given a set of articles and their operationalized features. Natural language processing (NLP) features (e.g., n-grams of part-of-speech tokens, etc.) can also be incorporated to enrich the input of the algorithms. After decisions are made, the system calculates its performance accuracy, providing another layer of analysis to decide if a piece of content can reliably be classified in a given category or not.

predictive power of the identified characteristics, and the combination of features that is optimal for detecting the eight types of content identified.

Note that we used the Decision Tree model to illustrate how various features that we have identified through our concept explication can be used to help machine learning models detect variants of "fake news." By following the root-to-leaf path in a Decision Tree, one can precisely "explain" why a particular article was determined to be a particular variant of "fake news" with respect to features (i.e., if-then-else rules;

Quinlan, 1986). In real settings, a more advanced but complicated supervised learning model, such as the Random Forest model, is likely to outperform the Decision Tree model in its classification accuracy (Ho, 1998). One can also use state-of-the-art Deep Learning models that are likely to improve the classification accuracy even more, but they run the risk of not being able to explain "why" it classified a piece of content in a certain category (Ivakhnenko & Lapa, 1973; LeCun et al., 1989). Such black-box deep learning models cannot explain why they make certain internal choices across layers of the network in the way that they did. While its output accuracy often outperforms popular supervised learning models significantly, its inability to satisfactorily explain "why" or "through what process" an article is classified to be a particular variant of "fake news" may lead users or applications to prefer supervised learning models, where the choice of features from our concept explication is very important for understanding the logic of classification.

The proposed list of features can additionally be utilized to develop literacy programs where users are provided with the eight types of content proposed in our taxonomy and their delineating characteristics. These programs will serve as toolkits to use when consuming online information, allowing users to value a piece of content for what it is and not confuse an advertisement or opinion piece, for example, with real news. Identifying and labeling content as belonging to a specific category will additionally help reduce reactance that is known to occur when readers are told a piece of content is blatantly false. This is because our classification provides a more nuanced and comprehensive differentiation among types of content encountered online. For example, in Nyhan and Reifler (2010), when participants were given attitude-discrepant corrective arguments, there was a backfire effect, strengthening their belief rather than changing it. Similar effects were found in Hart and Nisbet's (2011) research investigating polarization on controversial science issues. In these studies, participants assumed that the information they are reading is news. The reality is that Americans are no better than chance at distinguishing between facts and opinion statements (Mitchell, Gottfried, Barthel, & Sumida, 2018), nor can they differentiate news from promotional content (Stanford History Education Group, 2016). Differentiating between these types of content was easier before the Internet. For example, opinion was a section of its own in the newspaper, and different design elements were used to signal readers that a piece of content was an opinion rather than fact (Schallom, 2018). Similarly, with promotional content, banner ads used in traditional advertising are more likely to be differentiated due to their design, and in fact are often avoided by users (Drèze & Hussherr, 2003). This distinction between different types of content is blurred in the current information environment where information, satire, polarized content, opinion, promotional content, and so on, are all displayed in similar formats, and the content creator can be anyone with access to the Internet. Schallom (2018) suggests that guidelines should be established at an industry level to help users differentiate between types of content. Our taxonomy will help inform these guidelines while also providing fodder for literacy programs where we not only teach people about false and real information but also provide strategies to pinpoint key differentiating elements among the different types of content.

Discussion

This article has presented a first-cut explication of the concept of false news, with the goal of uncovering its different theoretical and operational definitions. This exercise, in turn, helped identify features to aid in detecting false news through a machine-learning algorithm. Our analysis unraveled various disagreements regarding the definition of false news and what content should and should not be included. This disagreement includes issues of intent, the inclusion or exclusion of satire content, and the nature of false news as binary or a continuum.

We propose conceptualizing false news such that any content, regardless of genre (political, sports, entertainment, etc.) and intentionality, can be assessed based on characteristics of its message, structure, sources, and network. Based on our explication, we identify a taxonomy of online content with unique features that can eventually be fed into a machine-learning algorithm for false news detection. In our taxonomy, we propose an analysis where no characteristic is the sole definer for classifying a piece of content as real/false news, but rather one of many indicators. Even though the list of features is not exhaustive, it can be used as a guideline for a multipronged analysis of online content. Due to the complexity of our online environment where anyone can create and disseminate content, the boundaries between types of content are increasingly blurred. The probabilistic functioning of machine learning provides a unique tool to combat the scourge of "fake news" because it assesses content based on a series of characteristics and learned rules, thus performing mutual disambiguation for estimating the probability that it is real, false, or any one of the other subcategories of "fake news" identified in our explication. Nevertheless, this area of study is rather nascent. Features or indicators of each type of content need to be further fine-tuned, developed, and tested.

Limitations of this study include the narrow search mechanism utilized for data collection. Since our interest was in identifying conceptual and operational definitions of false news, our search terms were limited to those referring to misinformation and its associated terms. It is possible that there are other articles exploring characteristics of the different types of content we identified that we did not account for. Second, in our research we focus on false news or made-up information with the purpose of making money or creating political division. We acknowledge the concrete focus might be a limitation when considering the broad, contested, and observer-dependent "fake news" phenomenon (Giglietto et al., 2016). Nevertheless, this concrete focus is imperative when utilizing machine learning methods of detection of such problematic content before it becomes viral. Furthermore, although we identified features of the message, source, structure, and network, the meaning analysis we performed outlined more characteristics of content than of the other three categories. As Lazer et al. (2018) explain, process and intent are arguably more important than content when identifying real versus false news because false information typically mimics the form of real information. Finally, as aforementioned, the identification of false news through machine learning is in its initial stages of development. It is important to flesh out the features to even greater level of detail than what we have specified, and test them.

Such limitations notwithstanding, this article has elucidated how the social sciences can complement computer sciences by informing the development of a machine-learning algorithm for "fake news" detection. Our research goal took the initial step to explicate "fake news" with the explicit purpose of identifying features that could be useful for machine learning. However, future research should computationally test which features work and which do not work for these purposes. For instance, it is possible that the current indicators can distinguish between certain types of content better than others. Take for example, misreporting. With the current indicators, it is likely that the algorithm could identify it once a news has been labeled as corrected, but this would be more difficult if this was not the case. Additionally, it is also important to test these features in different contexts and genres. For instance, while native advertisement might not be as prominent for political information, polarized content might be of greater importance. We also call for more content analyses of existing content belonging to each of our categories so that we can better understand their respective distinguishing features. This is especially true for citizen journalism, and promotional content, where our meaning analysis revealed fewer characteristics. We also suggest that these content analyses focus not only on content characteristics, but those of the network, source/intent, and structure.

In conclusion, we hope that the taxonomy proposed in our tables provide the foundation for identifying the major features that can aid reliable classification of "fake" and real news by both machines and humans, and thereby promote media literacy as well as a more credible information environment.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research is supported by the U. S. National Science Foundation (NSF) via Standard Grant No. CNS-1742702.

Notes

1. Retrieved from https://www.bbc.com/news/world-latin-america-47239060?intlink_from_url=https://www.bbc.com/news/topics/cp3mvpm3933t/venezuela-crisis&link_location=live-reporting-story
2. Retrieved from https://www.dcclothesline.com/2017/08/30/houston-looters-shoot-volunteer-rescuers-target-whites-and-trump-supporters/
3. Retrieved from https://www.snopes.com/news/2017/09/01/harvey-looting-troll-tweets/
4. Retrieved from https://www.nytimes.com/2019/03/07/opinion/ilhan-omar-anti-semitism.html
5. Retrieved from https://www.nytimes.com/interactive/2018/02/23/business/how-to-fill-out-1040-form.html
6. Retrieved from http://deepleftfield.info/trump-just-added-a-witness-tampering-charge-to-his-growing-list-of-crimes/

7. Retrieved from https://politics.theonion.com/poll-finds-100-of-americans-blame-shutdown-entirely-on-1831845310
8. Retrieved from https://www.theguardian.com/advertiser-content/stand-together/this-step-by-step-roadmap-to-end-homelessness

References

Adweek. (2009). *Landmark moments in citizen journalism*. Retrieved from https://www.adweek.com/digital/landmark-moments-in-citizen-journalism/

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*, 211-236. doi:10.1257/jep.31.2.211

American Press Institute. (2017). *What is journalism*. Retrieved from https://www.americanpressinstitute.org/journalism-essentials/what-is-journalism/

Argamon-Engelson, S., Koppel, M., & Avneri, G. (1998). *Style-based text categorization: What newspaper am I reading* (AAAI Technical Report WS-98-05). Retrieved from https://www.aaai.org/Papers/Workshops/1998/WS-98-05/WS98-05-001.pdf

Balmas, M. (2012). When fake news becomes real: Combined exposure to multiple news sources and political attitudes of inefficacy, alienation, and cynicism. *Communication Research*, *41*, 430-454. doi:10.1177/0093650212453600

Baym, G. (2005). The daily show: Discursive integration and the reinvention of political journalism. *Political Communication*, *22*, 259-276. doi:10.1080/10584600591006492

Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. New York, NY: Oxford University Press.

Berger, J. M. (2017). *The methodology of the Hamilton 68 dashboard*. Retrieved from http://securingdemocracy.gmfus.org/publications/methodology-hamilton-68-dashboard

Berghel, H. (2017). *Alt-news and post-truth in the "fake news" era*. Retrieved from https://ieeexplore.ieee.org/document/7912247

Bode, L., & Vraga, E. K. (2015). In related news, that was wrong: The correction of misinformation through related stories functionality in social media. *Journal of Communication*, *65*, 619-638. doi:10.1111/jcom.12166

Borden, S., & Tew, C. (2007). The role of journalist and the performance of journalism: Ethical lessons from "fake news" (seriously). *Journal of Mass Media Ethics*, *22*, 300-314.

Campbell, W. J. (2001). *Yellow journalism: Puncturing the myths, defining the legacies*. Westport, CT: Praeger.

Carlson, M. (2014). When news sites go native: Redefining the advertising–editorial divide in response to native advertising. *Journalism*, *16*, 849-865. doi:10.1177/1464884914545441

Chaffee, S. (1991). *Explication*. Newbury Park, CA: Sage.

Cohen, M. (2017). Fake news and manipulated data, the new GDPR, and the future of information. *Business Information Review*, *34*, 81-85. doi:10.1177/0266382117711328

Conroy, N. J., Rubin, V. L., & Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology*, *52*, 1-4. doi:10.1002/pra2.2015.145052010082

Cook, G. (2001). *The discourse of advertising* (2nd ed.). London, England: Routledge.

Cortes, C., & Vapnik, V. N. (1995). Support-vector networks. *Machine Learning*, *20*, 273-297.

Crain, R. (2017, February 6). Fake news could make advertising more believable. *Advertising Age*. Retrieved from https://adage.com/article/rance-crain/fake-news-make-advertising-believable/307836

Crocco, M., Halvorsen, A., Jacobsen, R., & Segall, A. (2017). Teaching with evidence. *Phi Delta Kappan*, *98*(7), 67-71. doi:10.1177/0031721717702635

Davis, W. (2016, December 5). Fake or real? How to self-check the news and get the facts: All tech considered. *NPR*. Retrieved from http://www.npr.org/sections/alltechconsidered/2016/12/05/503581220/fake-or-real-how-to-self-check-the-news-and-get-thefacts?utm_campaign=storyshare&utm_source=facebook.com&utm_medium=social

Downie, L., Jr., & Schudson, M. (2009). *The reconstruction of American journalism*. Retrieved from http://www.cjr.org/reconstruction/the_reconstruction_of_american.php

Dempsey, K. (2017, May). What's behind fake news and what you can do about it. *Information Today*. Retrieved from http://www.infotoday.com/IT/may17/Dempsey–Whats-Behind-Fake-News-and-What-You-Can-Do-About-It.shtml

Digital Resource Center. (2017). *Lesson 5: News vs. opinion*. Retrieved from http://drc.centerfornewsliteracy.org/content/lesson-5-news-vs-opinion

Drèze, X., & Hussherr, F. (2003). Internet advertising: Is anybody watching? *Journal of Interactive Marketing*, *17*(4), 8-23. doi:10.1002/dir.10063

Dylko, I., Dolgov, I., Hoffman, W., Eckhart, N., Molina, M., & Aaziz, O. (2017). The dark side of technology: An experimental investigation of the influence of customizability technology on online political selective exposure. *Computers in Human Behavior*, *73*, 181-190. doi:10.1016/j.chb.2017.03.031

Ecker, U. K. H., Lewandowsky, S., Chang, E. P., & Pillai, R. (2014). The effects of subtle misinformation in news headlines. *Journals of Experimental Psychology*, *20*, 323-335. doi:10.1037/xap0000028

Fake News. (2018). In *Cambridge online dictionary*. Retrieved from https://dictionary.cambridge.org/us/dictionary/english/fake-news

Fetzer, J. H. (2004). Disinformation: The use of false information. *Minds and Machines*, *14*, 231-240. doi:10.1023/B:MIND.0000021683.28604.5b

Frank, R. (2015). Caveat lector: Fake news as folklore. *Journal of American Folklore*, *128*, 315-332.

Funke, D. (2017, November 30). A satirical fake news site apologized for making a story too real. *Poynter*. Retrieved from https://www.poynter.org/news/satirical-fake-news-site-apologized-making-story-too-real

Giglietto, F., Iannelli, L., Rossi, L., & Valeriani, A. (2016). *Fakes, news and the election: A new taxonomy for the study of misleading information within the hybrid media system*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2878774

Hart, P. S., & Nisbet, E. C. (2011). Boomerang effects in science communication: How motivated reasoning and identity cues amplify opinion polarization about climate mitigation policies. *Communication Research*, *39*, 701-723. doi:10.1177/0093650211416646

Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*, 832-844. doi:10.1109/34.709601

Howard, P. N., Kollanyi, B., Bradshaw, S., & Neudert, L. M. (2017). *Social media, news and political information during the US election: Was polarizing content concentrated on swing states*. Retrieved from http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/09/Polarizing-Content-and-Swing-States.pdf

iReach. (2018). *Writing and formatting tips for news releases*. Retrieved from https://ireach.prnewswire.com/tips.aspx

Ivakhnenko, A. G., & Lapa, V. G. (1973). *Cybernetic predicting devices*. New York, NY: CCM Information Corporation.

Jack, C. (2017). *Lexicon of lies : Terms for problematic information. Data & Society Publication*. Retrieved from https://datasociety.net/pubs/oh/DataAndSociety_LexiconofLies.pdf

Jones, W. E. (2002). Dissident versus loyalist: Which scientists should we trust? *Journal of Value Inquiry*, *36*, 511-520. doi:10.1023/A:1021945707032

Klein, D. O., & Wueller, J. R. (2017). Fake news: A legal perspective. *Journal of Internet Law*, *20*(10), 1-13.

Kotsiantis, S. B. (2007). Supervised machine learning: A review of classification techniques. *Informatica*, *31*, 249-268.

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., . . . Schudson, M. (2018). The science of fake news. *Science*, *359*, 1094-1096.

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, *1*, 541-551. doi:10.1162/neco.1989.1.4.541

Maron, M. E. (1961). Automatic indexing: An experimental inquiry. *Journal of the ACM*, *8*, 404-417.

McCombs, M. (2005). A look at agenda-setting: Past, present and future. *Journalism Studies*, *6*(4), 543-557. doi:10.1080/14616700500250438

McElreath, R., & Smaldino, P. E. (2015). *Replication, communication, and the population dynamics of scientific discovery*. Retrieved from https://arxiv.org/abs/1503.02780

Mihailidis, P., & Viotty, S. (2017). Spreadable spectacle in digital culture: Civic expression, fake news, and the role of media literacies in "post-fact" society. *American Behavioral Scientist*, *61*, 441-454. doi:10.1177/0002764217701217

Mitchell, A., Gottfried, J., Barthel, M., & Sumida, N. (2018). *Distinguishing between factual and opinion statements in the news*. Retrieved from http://www.journalism.org/2018/06/18/distinguishing-between-factual-and-opinion-statements-in-the-news/

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. Cambridge: MIT Press.

Najmabadi, S. (2016, December 12). How can students be taught to detect fake news and dubious claims? *The Chronicle of Higher Education*. Retrieved from https://www.chronicle.com/article/How-Can-Students-Be-Taught-to/238652

Ng, Y., & Zhao, X. (2018). The human alarm system for sensational news, online news headlines, and associated generic digital footprints: A uses and gratifications approach. *Communication Research*. Advance online publication. doi:10.1177/0093650218793739

Nielsen, R. K., & Graves, L. (2017). *"News you don't believe": Audience perspectives on fake news*. Retrieved from https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2017-10/Nielsen&Graves_factsheet_1710v3_FINAL_download.pdf

Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, *32*, 303-330. doi:10.1007/s11109-010-9112-2

Ojala, M. (2017). Fake business news. *Online Searcher*, *41*(3), 60-62.

Padgett, L. (2017). Filtering out fake news: It all starts with media literacy. *Information Today*, *34*(1), 6.

Pariser, E. (2011). *The filter bubble: How the new personalized Web is changing what we read and how we think*. New York, NY: Penguin Press.

Pennycook, G., & Rand, D. G. (2017). *Who falls for fake news? The roles of analytical thinking, motivated reasoning, political ideology, and bullshit respectively*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3023545

Potthast, M., Kiesel, J., Reinartz, K., Bevendorff, J., & Stein, B. (2017). *A stylometric inquiry into hyperpartisan and fake news*. Retrieved from https://arxiv.org/pdf/1702.05638.pdf

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, *1*(1), 81-106.

Reilly, I. (2012). Satirical fake news and/as American political discourse. *Journal of American Culture*, *35*, 258-275.

Reilly, I. (2013). From critique to mobilization: The yes men and the utopian politics of satirical fake news. *International Journal of Communication*, *7*, 1243-1264.

Rich, P. R., & Zaragoza, M. S. (2016). The continued influence of implied and explicitly stated misinformation in news reports. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *4*, 62-74. doi:10.1037/xlm0000155

Roozenbeek, J., & van der Linden, S. (2019). The fake news game: Actively inoculating against the risk of misinformation, *Journal of Risk Research*, *22*, 570-580. doi:10.1080/13669877.2018.1443491

Samuel, A. (2016). *To fix fake news, look to yellow journalism*. Retrieved from https://daily.jstor.org/to-fix-fake-news-look-to-yellow-journalism/

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., . . . Jinks, C. (2018). Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality & Quantity*, *52*, 1893-1907. doi:10.1007/s11135-017-0574-8

Schallom, R. (2018). *Better design helps differentiate opinion and news*. Retrieved from http://www.niemanlab.org/2017/12/better-design-helps-differentiate-opinion-and-news/

Shao, C., Ciampaglia, G. L., Varol, O., Yang, K. C., Flammini, A., & Menczer, F. (2018). The spread of low-credibility content by social bots. *Nature Communications*, *9*, 1-9. doi:10.1038/s41467-018-06930-7

Sheble, L. (2018). Misinformation and science: Emergence, diffusion, and persistence. In B. G. Southwell, E. A. Thorson, & L. Sheble (Eds.), *Misinformation and mass audiences* (pp. 157-176). Austin: University of Texas Press.

Shoemaker, P. J. (2017). News values: Reciprocal effects on journalists and journalism. In P. Rössler, C. A. Hoffner, & L. van Zoonen (Eds.), *The international encyclopedia of media effects* (pp. 1-9). New York, NY: Wiley. doi:10.1002/9781118783764.wbieme0053

Shu, K., Wang, S., Le, T., Lee, D., & Liu, H. (2018). Deep headline generation for clickbait detection. In *2018 IEEE International Conference on Data Mining (ICDM)*, 467-476.

Silverman, C. (2016, November 16). This analysis shows how viral fake election news stories outperformed real news on Facebook. *BuzzFeed*. Retrieved from https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook?utm_term=.ilJOyKYld#.lyk3Kb9NJ

Soares, I. (2017). *The fake news machine: Inside a town gearing up for 2020*. Retrieved from http://money.cnn.com/interactive/media/the-macedonia-story/

Southwell, B. G., Thorson, E. A., & Sheble, L. (2017). The persistence and peril of misinformation. *American Scientist*, *105*, 372-375.

Stanford History Education Group. (2016). *Evaluating information: The cornerstone of civic online reasoning*. Retrieved from https://sheg.stanford.edu/upload/V3LessonPlans/Executive%20Summary%2011.21.16.pdf

Storyzy. (2017). *About*. Retrieved from https://storyzy.com/about

Sundar, S. S., & Nass, C. (2001). Conceptualizing sources in online news. *Journal of Communication*, *51*, 52-72. doi:10.1111/j.1460-2466.2001.tb02872.x

Sunstein, C. R. (2002). The law of group polarization. *Journal of Political Philosophy*, *10*, 175-195. doi:10.1111/1467-9760.00148

Tandoc, E. C., Lim, Z, W., & Ling, R. (2018). Defining "fake news." *Digital Journalism*, *6*, 137-153. doi:10.1080/21670811.2017.1360143

Trilling, D., Tolochko, P., & Burscher, B. (2017). From newsworthiness to shareworthiness: How to predict news sharing based on article characteristics. *Journalism & Mass Communication Quarterly*, *94*, 38-60. doi:10.1177/1077699016654682

Turner, M. (2017, July 23). *Commentary vs. journalism: Are journalists biased?* Retrieved from https://www.schooljournalism.org/commentary-vs-journalism-are-journalists-biased/

Vorhies, W. (2017, May 1). *Using algorithms to detect fake news: The state of the art*. Retrieved from https://www.datasciencecentral.com/profiles/blogs/using-algorithms-to-detect-fake-news-the-state-of-the-art?

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*, 1146-1151. doi:10.1126/science.aap9559

Waisbord, S. (2018). Truth is what happens to news. *Journalism Studies*, *19*, 1866-1878. doi:10.1080/1461670X.2018.1492881

Wang, W. Y. (2017). *"Liar, liar pants on fire": A new benchmark dataset for fake news detection*. Retrieved from https://arxiv.org/abs/1705.00648

Wardle, C. (2017, February 16). *Fake new: It's complicated*. Retrieved from https://firstdraft-news.org/fake-news-complicated/

Yarros, V. S. (1922). Journalism, ethics, and common sense. *International Journal of Ethics*, *32*, 410-419.

Author Biographies

Maria D. Molina is a doctoral student (ABD) in the Donald P. Bellisario College of Communications at The Pennsylvania State University. Maria's research builds upon psychology and communication theory to advance interface design and user experience of digital media technologies. Specifically, her research focuses on exploring how and why people share online and the role of technological affordances at promoting sharing behavior.

S. Shyam Sundar (http://bellisario.psu.edu/people/individual/s.-shyam-sundar) is James P. Jimirro professor of Media Effects and founding director of the Media Effects Research Laboratory at The Pennsylvania State University. His research investigates social and psychological effects of technological affordances in interactive, digital media. He edited the first-ever *Handbook on the Psychology of Communication Technology* (Blackwell Wiley, 2015) and served as editor-in-chief of the *Journal of Computer-Mediated Communication* from 2013 to 2017.

Thai Le is pursuing his Ph.D. degree at The Pennsylvania State Univ ersity. He is majoring in Information Science and Technology. Thai mainly contributes in the field of Machine Learning and AI. His research lies in the area of interpretable AI, large-scale synthetic data generation and augmentation.

Dongwon Lee is an associate professor in the College of Information Sciences and Technology (a.k.a. iSchool) of The Pennsylvania State University, USA. He researches broadly in Data Science on the management and mining of data in diverse forms including structured records, texts, graphs, social media, and Web. Further details of his research can be found at: http://pike.psu.edu/.