American Journal of Art and Communication
Vol.7, Issue 2; March-April - 2022;
1252 Columbia Rd NW, Washington DC, United States
https://topjournals.org/index.php/AJAC; mail: topacademicjournals@gmail.com

THE PROMISE AND CHALLENGES OF AUTOMATED FACT-CHECKING FOR PROFESSIONAL PRACTICES

Kari Anne Olsen, Stian Andersen, Kristine Berg
University of Bergen, Norway

Abstract: The prevalence of disinformation, misinformation, and propaganda online has been on the rise, with multiple political, societal, and cultural implications. Automated fact-checking (AFC) has increasingly garnered interest in research over the past years as a way to address this problem. AFC systems use natural language processing and machine learning techniques to identify and verify claims made online. While AFC systems have shown promise in some areas, they still face challenges, such as the ability to handle complex claims and the need for human supervision.

Keywords: Automated fact-checking (AFC)**,** Disinformation**,** Misinformation**,** Propaganda**,** Natural language processing (NLP)

Introduction

The prevalence of disinformation, misinformation, and propaganda online has been on the rise, accompanied by multiple political, societal, and cultural implications (Kalsnes, 2018). As such, automated fact-checking (AFC) has increasingly garnered interest in research over the past years (Guo et al., 2022). These different forms of information disorder are not new to the world's history (Kalsnes, 2018; Posetti & Matthews, 2018), but online social interactions contribute to their large-scale accessibility and visibility, considering the central role played by social media and platforms (Guess et al., 2020).

Even though the verification of information has always been a part of editorial processes, it was popularized as a prominent subgenre in recent years (Singer, 2021), referring to "good journalism" (Singer, 2018) or "unbiased journalism" (Graves, 2018). At the same time, the fact-checking community also encompasses political activists or fact-checkers acting on behalf of social progress (Mena, 2019). Although they share the same idea of defending and practicing a form of accountability reporting, these actors may view this ideal differently (Graves, 2018). Such heterogeneity is challenging for AFC tool designers, particularly in defining the requirements or needs of potential end users whose various organizational contexts and professional standards contribute to shaping their practices.

In this article, we define the end user as a fact-checker employed either in a fact-checking organization in a newsroom model or in a newsroom alongside or not alongside other journalistic activities. These professionals assimilate their identity and professional practices into journalism, considering that factchecking is a genre of journalism (Cavaliere, 2020; Graves & Cherubini, 2016). Therefore, the professional values that frame their practices are linked to the traditional ethical values of journalism—accuracy, fairness, and objectivity (Frost, 2015)—that support the social responsibility of the journalists regarding the audiences to whom they are accountable (Bardoel & d'Haenens, 2004). These principles are strongly connected to the transparency defended in the ethical principles of the International Fact-Checkers Network (IFCN; Mena, 2019) and the European Fact-Checking Standards Network, which bring together fact-checkers to promote excellence in fact-checking.

Considering the speed at which misinformation and disinformation spread, AFC can be fast and effective tools not only to find claims worth fact-checking, relevant previously fact-checked claims, or supporting evidence but also to translate or summarize content (Nakov et al., 2021). The four stages of manual fact-checking are extracting statements, constructing appropriate questions, obtaining answers from relevant sources, and reaching a verdict using these answers (Vlachos & Riedel, 2014). Concretely, they encompass four primary tasks: monitoring media and capturing content, detecting claims, checking claims, and publishing content (Konstantinovskiy, Price, Babakar, & Zubiaga, 2021). These activities require the conjugation of critical thinking and know-how related to online investigative journalism and multimedia forensics.

The boundary object (BO) theory allows us to approach AFC tools as technological artifacts that exist between adjacent communities of designers and end users. According to this theory, as boundary objects, AFC tools "negotiate meaning to understand and articulate connections and disconnections between communities, cultures, and information infrastructures" (Huvila, Anderson, Jansen, Mckenzie, & Worrall, 2017, p. 1812). The BO theory can be seen as a valuable instrument to highlight the issues that researchers face when technology and journalism intertwine because it involves distinct occupational and social worlds that do not necessarily share the same values or visions of journalism (Lewis & Usher, 2016; Sirén-Heikel, Kjellman, & Lindén, 2022). From the point of view of the approach to the relationship between design and use (or between developers and fact-checkers), these issues are potentially linked to the work of interpreting what fact-checking is, that is, in terms of objectives or in terms of process. Therefore, it would involve more than an accurate understanding or transfer of knowledge from fact-checkers to designers because interdisciplinary interactions transcend differences in meaning between disciplines (Akkerman & Bakker, 2011; Fox, 2011; Trompette & Vinck, 2009).

This systematic literature review maps the field of AFC and its progress over the past five years to follow the recent developments in artificial intelligence (AI) in this field. It identifies how AFC tools, as boundary objects, connect or disconnect adjacent communities on either side of the technological artifact. In other words, how do they adapt to journalism and infuse professional practices? The objective is to provide a comprehensive multidisciplinary state of the art that considers a holistic and sociotechnical approach to studying AFC from a journalistic perspective to nourish future works. Therefore, two research questions are investigated:

RQ1: To what extent does research on AFC tools consider end users, and how?

RQ2: As boundary objects, how do AFC tools connect and disconnect with the social world of journalism?

Method

A systematic literature review is a means of collecting and synthesizing previous research, providing an overview of areas covered by the research, and demonstrating evidence on a meta-level (Snyder, 2019). Its main objective is to answer specific questions by relying on rigorous and explicit methods to ensure the work's transparency, transferability, and replicability (Thomas & Harden, 2008). Although systematic literature reviews are commonly used in the medical and computer sciences, they can also be utilized in social sciences to provide an overall picture of the evidence in a topic area to guide future research (Petticrew & Robert, 2006, p. 21). Meta-analysis refers to statistical techniques to obtain overall estimations or a synthesis of published research works (Çoğaltay & Karadağ, 2015; Shelby & Vaske, 2008). In this study, the meta-analysis was mobilized through descriptive statistics to highlight the corpus characteristics and provide quantified evidence to complement the systematic literature review (Davis, Mengersen, Bennett, & Mazerolle, 2014; Mengist, Soromessa, & Legese, 2020).

From a methodological perspective, this systematic literature review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to ensure transparent and complete reporting (Liberati et al., 2009). It consists of a checklist of 27 items that frame the method, the writing of a systematic review report, and a flow diagram that shapes information retrieval and selection (Page et al., 2021). The tool Parsifal, available online (Durier da Silva, Bicharra Garcia, & Matsui Siqueira, 2022), was used to define the research planning (objectives, research questions, keywords, search strings, sources, and selection criteria) as well as to identify the articles published in Scopus. The tool was also used to import all the collected data as spreadsheets containing the articles' metadata to identify duplicate entries. The data collection was carried out via Google Scholar, Semantic Scholar, and Scopus between March 14 and 31, 2022. The search terms were approached through three complementary queries to refine the quality of the corpus and limit the phenomena of noise (linked to unexpected results) and silence (referring to the absence of expected results). This iterative process followed an inverted pyramid model, that is, going from the general to the specific, considering that Google Scholar was the most general database used and Scopus was the most specific:

- Fact-checking AND "machine learning" (Google Scholar)
- Automated AND fact-checking AND journalism (Semantic Scholar)
- ("machine learning" OR automated) AND fact-checking (Scopus)

These three queries returned 918 results. Inclusion and exclusion criteria were applied to refine the corpus selection. Published peer-reviewed articles; open-access articles published on arXiv, which are published after moderation but are not peer-reviewed; book chapters; and proceedings related to AFC were included because the first objective of this review was to get a broad overview of the research works. Duplicate entries were removed along with all the articles that fell into the scope of the exclusion criteria: non-English texts, articles dated before 2017, undated articles, articles unrelated to AFC, and articles not available either in a PDF format or online. Dissertations were excluded from the review due to their extensive length, which is not comparable to research articles, and their limited accessibility and availability. Moreover, because their authors could publish their work in peer-reviewed journals, this exclusion criterion was also intended to limit the number of duplicate research. After the processes, the main corpus consisted of 267 articles used for a state of the art (Figure 1).

Retrieving (918)		
Google Scholar: 205	Semantic Scholar: 338	Scopus: 375
Type of paper + close and distant reading (267, main corpus)		
Google Scholar: 101	Semantic Scholar: 94	Scopus: 72
Users OR journalism (72, subcorpus 1)		
Google Scholar: 42	Semantic Scholar: 18	Scopus: 12
Users AND journalism (21, subcorpus 2)		
Google Scholar: 11	Semantic Scholar: 3	Scopus: 7

Figure 1. Corpus selection process.

To fit the scope of this study, the selection criteria were addressed using two complementary methods: (1) close reading of all of the collected abstracts and, in cases of uncertainty, additional close reading of the full texts, which concerned half of the collected corpus, and (2) distant reading of the abstracts through text mining and text analysis techniques using the programming language R (Ramage, Rosen, Chuang, Manning, & Mcfarland, 2009; Welbers, Van Atteveldt, & Benoit, 2017) and dedicated packages to proceed a meta-analysis and n-grams frequencies (tidyverse, tidytext, TM, quanteda, highcharter), topic modeling (LDA) and clustering (textmineR). This process allowed us to define a first subcorpus of 72 articles that considered users or journalism, as well as a second subcorpus of 21 articles that referred to users and journalism (Figure 1).

The R packages previously mentioned were also used for the meta-analysis of the main corpus. The meta-analysis combined a deductive and an inductive approach (Grimmer, Roberts, & Stewart, 2021; Molina & Garip, 2019) to support discoveries considering research questions that globally refer to the challenges of AFC in the social world of journalism. In addition, we have created a database including a unique identifier, the title of the article, the abstract, the field of research, the type of article, the year of publication, and the number of citations. We added two columns to identify articles related to users and journalism or newsrooms. The datasets and the source code of all these operations are available on GitHub (https://github.com/laurence001/AFC_SLR).

The limitations of the research strategy are related to the level of accuracy provided by distant reading. Indeed, it is recognized that topic modeling is not suitable for advanced data relationships and performs poorly when documents do not have a sufficient length (Vayansky & Kumar, 2020). Clustering is also challenging for finding similarities between data points and grouping similar ones into the same cluster (Qaddoura, Faris, & Aljarah, 2020). Consequently, these results were primarily used to support human analysis. In addition, it is challenging to claim the completeness of a corpus queried through databases during a short period. Nevertheless, the examination of the references in the articles collected indicates the representativeness of the corpus on recent research works on AFC.

Results

The presentation of the results is divided into three parts: (1) a meta-analysis of the corpus to identify the state of the art of AFC technologies as well as the general challenges of AFC in journalism, (2) a review of the first subcorpus related to the uses and users of AFC tools to answer RQ1, and (3) an analysis of the second subcorpus identifying the challenges of developing AFC tools in a journalistic context to address RQ2. This three-level analysis aims to provide evidence to fuel the discussion about how research should consider end users more when designing AFC tools.

Meta-Analysis and Challenges for Journalism

Most of the articles collected were published in 2019, 2020, and 2021 (respectively 22.47%, 22.10%, and 29.96%). However, it does not mean that AFC gained particular interest over this period, considering the time required for the research and the time needed for reviewing and publishing, although preprint articles represent 20.52% of the main corpus. This corpus also includes 51.49% of articles, 46.64% of proceedings, and 1.87% of book chapters. Computer science was the main research area covered (80.90%), followed by social computing (9.74%), and the articles related to social science or journalism studies represented less than 7% of the corpus. It is not surprising that computer science is the most represented field given that the development of AFC tools involves using specific technologies such as machine learning, natural language processing, and knowledge graphs (Gallofré Ocaña & Opdahl, 2020; Lakshmanan, Simpson, & Thirumuruganathan, 2019). In addition, social computing and information science are closely related to computer science, and journalism studies and social science intersect.

AFC covers four main functionalities that assist and support human fact-checkers: finding claims, detecting already fact-checked claims, evidence retrieval, and verification (Nakov et al., 2021). Based on this typology, automated analysis of the abstracts of the main corpus demonstrated that the current state of the art in AFC technologies primarily covers claim detection (34.33%), followed by claim verification (14.18%) and evidence retrieval (8.20%). The technologies developed in AFC are most frequently oriented to machine learning (supervised and unsupervised), followed by natural language processing and knowledge graphs. Blockchain technologies are marginally represented (0.75%).

AFC technologies are essentially text focused. Only 14 articles were dedicated to images and/or videos (5.24%). They included datasets of images and/or videos (e.g., Papadopoulou, Zampoglou, Papadopoulos, & Kompatsiaris, 2019; Zlatkova, Nakov, & Koychev, 2019), supervised machine learning to detect deceptive images or for image classification (Boididou et al., 2018; Reis & Benevenuto, 2021), the assessment of image forensics services (Katsaounidou, Gardikiotis, Tsipas, & Dimoulas, 2020; Nakov et al., 2021), social-computing solutions for user-generated content verification (Middleton, Papadopoulos, & Kompatsiaris, 2018), and deepfake detection (Hoque, Ferdous, Khan, & Tarkoma, 2021). Four articles focused on multimodal solutions to detect misleading texts and images (Abdelnabi, Hasan, & Fritz, 2021; Dhankar, Zaïane, & Bolduc, 2022; Gao, Hoffmann, Oikonomou, Kiskovski, & Bandhakavi, 2022; Yang et al., 2018). One article was dedicated to verifying audio content associated with video (Vryzas, Katsaounidou, Vrysis, Kotsakis, & Dimoulas, 2022). The full-text search results indicated that 129 articles were about experimental-stage systems (35%), and 179 articles were about frameworks (48.5%). In addition, the authors of 45 articles provide online the source code of their fact-checking system (60%) and/or the datasets they built (62.2%) through a GitHub project page.

Based on the abstracts' content, datasets are the most prominent topic because it was found in 40% of the corpus. This is not surprising because AFC tools rely on data for classification and prediction purposes: it is how machine-learning works. Several articles focused more specifically on creating training datasets that are specific to one particular news context (e.g., on the Syrian war, U.S. politics, or COVID19). They are mainly published in English, except for one multilingual article and three others in Spanish, Arabic, and Czech. They rely on several different classification systems such as "True," "False," or "Halftrue" (Wang, 2017); "Unproven" (Kotonya & Toni, 2020); and "Contradiction," "Compatible," or "Unrelated" (Sepúlveda-Torres, Bonet-Jover, & Saquete, 2021). They also encompass variable amounts of data; for instance, the FEVER dataset—which consists of 185,445 claims generated by altering sentences extracted from Wikipedia—comprises 185,000 rows (Thorne, Vlachos, Christodoulopoulos, & Mittal, 2018), whereas AraStance has more than 4,000 rows. It consists of a multicountry and multidomain dataset of Arabic stance detection for fact-checking, based on claim-article pairs from a diverse set of sources comprising three factchecking websites and one news website (Alhindi, Alabdulkarim, Alshehri, Abdul-Mageed, & Nakov, 2021).

The reliability of dataset labeling (or annotation) required for supervised tasks is challenged by crowdsourcing as mentioned in the abstracts of 11 articles. At the same time, Wikipedia was quoted as a source for data extraction in five out of 12 articles, and it was mentioned in 110 articles (41.2% of the corpus). Training journalistic tools with Wikipedia poses several issues because the content is generated by unknown users (Umarova & Mustafaraj, 2019), and encyclopedic texts differ from journalistic ones. From the broader perspective of machine-learning research, it has been pointed out that the crowd is not always made up of experts and that human biases can interfere with tasks' accuracy and outcomes' reliability (Lease, 2011; Miceli, Posada, & Yang, 2021).

The performance of the AFC systems relying on machine learning can be approached through the F1 score that is commonly used to evaluate the precision and recall of a classification model. Hence, it is tackled as an error rate indicator. We found 92 articles (35.20% of the corpus) referring to this score, which ranged between 5% (Hui Xian Ng & Carley, 2021) and 99.6% (Ebadi, Choo, & Rad, 2022). Although the F1 score is not the only indicator to measure the performance of a model, it remains generic and can be considered a weak estimator for the uncertainty that can remain in the outcomes (Kläs & Vollmer, 2018). The performance of a model can also be tackled through the detection of overfitting, which occurs when the model is trained for too long and, as a result, comes to reflect the specifics of the training data rather than the general characteristics of the underlying domain. Consequently, the model fails to find a general predictive rule (Dietterich, 1995). In this corpus, the overfitting of the machine-learning model was pointed out in 24 articles. This does not mean that the research has failed, but it illustrates the challenges in developing efficient and scalable AFC tools.

As boundary objects, AFC systems represent a clear division between their designers and factcheckers. This finding is confirmed in the second part of the analysis, which focuses on the relationship between AFC systems and the social world of journalism. However, some attempts to connect technologies to work environments and practices have also been observed, particularly when considering keeping the human fact-checker in the loop.

Uses and End Users of AFC Systems

The proportion of articles considering end users and/or journalism is relatively low, representing 20.22% and 14.61% of the corpus, respectively. This highlights that attempts by developers in AFC are driven by the goal to provide a technical solution to the social problem of information disorder. According to this approach, the growing importance of information disorder creates a need for new tools to speed up the process of detecting and verifying claims (e.g., Azevedo, 2018; Çarik & Yeniterzi, 2021; Du, Bosselut, & Manning, 2022; Kar, 2020). However, we found 72 articles directly or indirectly related to the sociotechnical dimension of AFC technologies, that is, 26% of the main corpus, comprising the first subcorpus devoted to the uses and end users.

Several articles highlighted the time-consuming aspects of human fact-checking activities and the challenge that human fact-checkers cannot keep up with the amount of misinformation and the speed at which it spreads (e.g., Adair, Li, Yang, & Yu, 2017; Gencheva, Koychev, Màrquez, Barrón-Cedeño, & Nakov, 2019; Hanselowski & Gurevych, 2017; Jiang, Baumgartner, Ittycheriah, & Yu, 2020). Scholars also emphasized that human fact-checking requires expertise (Jaroucheh, Alissa, Buchanan, & Liu, 2020) and, in a certain way, human intuition and creativity that can't be automated (Nakov et al., 2021). Furthermore, the relationship between AFC tools and human users remains problematic (Borges, Martins, & Calado, 2019) due to limitations related to credibility issues for automated systems and scalability issues for human factchecking (Nakov et al., 2021). When researchers consider uses or end users, the focus is mostly on technological solutions to speed up the AFC process or assist human fact-checkers (e.g., Brand, Roitero, Soprano, & Demartini, 2021; Majithia et al., 2019). However, the fact-checkers toolbox is also explored (Bañon Castellón, 2021; Nygren, Guath, Axelsson, & Frau-Meigs, 2021; Svahn & Perfumi, 2021), compared (Školkay & Filin, 2019), classified (Nakov et al., 2021), and evaluated (Komendantova et al., 2021; Picha Edwardsson, Al-Saqaf, & Nygren, 2021). Another perspective is related to the benefits of human-computer interactions (Miranda et al., 2019; Shi, Bhattacharya, Das, Lease, & Gwizdka, 2022; Yang et al., 2019). These approaches can be understood in terms of challenges and opportunities that concern both users and developers (e.g., Demartini, Mizzaro, & Spina, 2020; Gallofré Ocaña & Opdahl, 2020). They encourage a hybrid process where humans are augmented by the use of AI, broadening a technical solution into a sociotechnical one, which is likely to result in the reconfiguration of professional practices (Diakopoulos, Trielli, & Lee, 2021). This concerns not only the AFC system's functionalities but also how the tool was designed, considering the users' requirements (Nguyen et al., 2018) or the transparency of the model in facilitating human interactions (Rony, Hoque, & Hassan, 2020).

Making an AFC system transparent depends on its explicability. Transparency concerns both the upstream components of the tool (Reis, Correia, Murai, Veloso, & Benevenuto, 2019)—for example, the explanation of the model and the information sources used (Gencheva et al., 2019) or why the expert assigned that label to a training dataset (Berendt et al., 2021)—its interface (Katsaounidou et al., 2020); and the results it provides (Denaux & Gomez-Perez, 2020; Middleton et al., 2018). Therefore, the use of AI tools is first and foremost a matter of trust (Demartini et al., 2020). Providing explanations not only helps humans perform fact-checking but also can foster trust toward the tool or improve AFC systems (Gad-Elrab, Stepanova, Urbani, & Weikum, 2019). However, this can be difficult when the system relies on black-box methods, such as deep neural networks (Saeed & Papotti, 2021).

Contextual information about the application domain is another topic that has characterized this first subcorpus. It implies some shared expertise between coders and end users (Berendt et al., 2021). Providing context with the results contributes to a better understanding of a claim (Vizoso, Vaz-Álvarez, & López-García, 2021; Vryzas et al., 2022) and suggests trustworthiness in the tool and its results (Middleton et al., 2018). In addition, the lack of knowledge of the application domain may interfere with the correctness and the quality of a training dataset's annotations (Singh, Das, Li, & Lease, 2021). It was also pinpointed that human fact-checking also requires human judgment and sensitivity (Graves, 2018). Therefore, research in AFC also considers contextbased approaches (Boididou et al., 2018; Fairbanks, Fitch, Bradfield, & Briscoe, 2020; Gencheva et al., 2019), which can be interpreted as another form of an attempt to link boundary objects with the knowledge and practices of a given professional community (Trompette & Vinck, 2009).

AFC and Journalism

The analysis of the second subcorpus of 21 articles (7.86% of the main corpus) demonstrated that the most prominent perspective is about developing tools (42.85%) or supporting journalistic practices (52.38%), mainly for claims or stance detection (33.33%). Nearly half of the articles (42.86%) presented prototypes (e.g., Adair et al., 2017; Horne, Dron, Khedr, & Adali, 2018), of which five consisted of experiments (Adler & Boscaini-Gilroy, 2019; Boididou et al., 2018; Jiang et al., 2020; Masood & Aker, 2018; Vryzas et al., 2022). In addition, one article focused on fact-checkers' needs by exploring the tools they use and how AFC tools can support them (Nakov et al., 2021), and four articles—all related to journalism studies—were dedicated to journalism practices within newsrooms (Bañon Castellón, 2021; de Haan, Van Den Berg, Goutier, Kruikemeier, & Lecheler, 2022; Diakopoulos, 2020; Picha Edwardsson et al., 2021).

The world of journalism is less considered in research on AFC. However, when it is considered, the complementarity between the journalist and the tool is the most frequently underlined aspect. That is mainly because fact-checking is a time-consuming process that still requires human input either to assess the validity of a claim or to adapt to dynamic news contexts by extending the daily collection of the source, for instance, as news information collection is a continuous process that occurs in a moving reality (Berendt et al., 2021; Hassan et al., 2017). Several articles also stressed that human-machine complementarity is still needed because AFC systems are not strong enough to give rise to fully automated solutions (Komendantova et al., 2021).

The question of the fact-checkers' user needs was tackled either through a sociotechnical prism (Diakopoulos et al., 2021) or by considering the cultural background and attitudes of journalists who may be skeptical toward technology and who generally lack algorithmic culture (de Haan et al., 2022). Still, from a journalistic point of view, other challenges are related to the "lack of collaboration between researchers and practitioners in terms of defining tasks and developing datasets for automated systems" (Nakov et al., 2021, p. 5), the lack of time to learn how to use digital tools (Picha Edwardsson et al., 2021), the lack of consideration of the specific needs that can be expressed by fact-checkers, and the lack of transparency of some fact-checking services (Komendantova et al., 2021). Although addressing some of these issues might be seen as a fruitful path to better intersect AFC systems' users and designers, all cannot be technically solved, mainly because they also rely on social and organizational variables that involve human decisions.

Designing an AFC system is also challenging because of the complexity of defining central concepts such as "claim" and "verification" (e.g., Konstantinovskiy et al., 2021). Other issues can also be related to the limited availability of training datasets or to the overall quality of some existing ones, including the aspects related to the quality of their annotations (e.g., de Haan et al., 2022). Systems must also be adapted to the complexity of the newsroom's workflows, where the multiplication of information channels is required to speed up the processes and provide newsworthy insights (Gallofré Ocaña & Opdahl, 2020).

Discussion and Conclusion

Regarding RQ1, related to the integration of end-user perspectives, the findings indicated that most research focuses on providing technological solutions to a social problem without embedding end users' views or needs. As machine-learning systems rely on training data, quality issues appear crucial, specifically when annotations are based on crowdsourcing by nonexperts. Annotations are inherently prone to errors even when control procedures are set for correction (Northcutt, Athalye, & Mueller, 2021). They are also language, domain, and context dependent, making them less adaptive or reusable. In addition, the datasets used for machine learning have various sizes, rely on binary classes for the labeling, and tend to be unbalanced, which is not ideal for training (Zeng, Abumansour, & Zubiaga, 2021).

Because they are likely to rely on various annotation process types, the question is also that of developing a standardized annotation system (Zeng et al., 2021). Such diversity also complicates both comparing approaches and gauging their improvement over time. At the same time, there is no consensus on the best classification strategies and sets of features for AFC tools developed for detection (Silva, Santos, Almeida, & Pardo, 2020). Nonetheless, researchers have underlined that the datasets developed for natural language processing purposes are more valuable, cover different domains, and help progress research in automated claim validation (Zeng et al., 2021). Furthermore, automating detection and selection processes is also challenging because the criteria are likely to vary from one fact-checking organization to another, although common patterns are observed, such as prominence within the opinion or debate, virality, and measures of social engagement (Micallef, Armacost, Memon, & Patil, 2022).

Working with data from datasets based on Wikipedia raises the same risk of potential quality issues as any other user-generated content in terms of accuracy or reliability because there are no guarantees about the users' expertise. In this regard, quality control of the datasets and their maintenance over time also appeared as two other obstacles to efficiency, and data quality has "a huge impact on the efficiency, accuracy and complexity of machine learning tasks" (Gupta et al., 2021, p. 1). In addition, the use of crowd workers raises "practical and ethical issues, such as funding and remuneration" (Berendt et al., 2021, p. 11). Although tremendous efforts are dedicated to gathering and annotating datasets for machine-learning tasks, the lack of relevant ones could be a constraint on developing AFC tools (Pathak & Srihari, 2019).

At the same time, the lack of vision about the fact-checkers' needs and requirements appears problematic as several fact-checking tools are developed without considering them (Komendantova et al., 2021). The explicability of the system contributes to building a relationship of confidence between the user and the tool, and it plays a central role in using AI tools (Stray, 2019). In a context where fact-checkers vigorously promote values of transparency, being able to explain algorithms at work and the results they provide Katsaounidou et al., 2020) can be considered a lever for trust, or an instrument to demystify tools that might be considered black boxes (e.g., Bartneck, Lutge, Wagner, & Welsh, 2021; de Haan et al., 2022; Zhou, Hu, Li, Yu, & Chen, 2019).

From the BO theory perspective addressed in RQ2, AFC tools might exist between the adjacent communities of researchers and fact-checkers by involving the latter in the process when it comes to assessing the performance and the usability of the tool, for instance because this is likely to help improve it (Miranda et al., 2019). This implies a transfer of knowledge from one professional community to another, although the computational process results from a work of interpretation (Huvila et al., 2017). The social aspects of fact-checking also play a role in the adoption of the technological artifact, including in terms of its perceived advantages (Fox, 2011). Specific training in identifying manipulated images and deep fakes, along with the ease of use and integration of AFC tools into editorial systems, are additional outcomes from journalism studies to enhance the integration of AFC tools in newsrooms (Katsaounidou et al., 2020; Picha Edwardsson et al., 2021; Vizoso et al., 2021).

AFC tools are helpful to detect falsehoods, but they do not eliminate the need for human intervention (Bañon Castellón, 2021). Moreover, AFC tools should meet the users' trust (Nakov et al., 2021), which can be achieved only if both the system and its outcomes are trustworthy too. When professional fact-checkers use their human expertise and take charge of the labeling of training datasets for machine-learning-based systems (Berendt et al., 2021), this "human in the loop" perspective is a means to achieve connection between the communities surrounding the BO. Integrating the users' knowledge also improves a system's transparency and enhances its trustworthiness (Nguyen et al., 2018) because it also emphasizes on the social components of technology. Ultimately, "automated fact-checking works well in some cases," but it "still needs improvement prior to widespread use" (Lazarski, Al-Khassaweneh, & Scotts Howard, 2021, p. 1).

Reconnecting adjacent professional communities from either part of the BO may also be achieved by paying acute attention to the users' beliefs and the possibility "to infuse their views and knowledge into the system" (Shi et al., 2022, p. 315). Furthermore, the necessary compatibility of a technological artifact with journalistic ideals and values was also stressed in the context of the diffusion of news automation within newsrooms (Diakopoulos, 2019). However, in journalism, psychological and cultural barriers challenge the effectiveness of AFC tools, although technology can make fact-checking easier and faster (Cazalens, Lamarre, Leblay, Manolescu, & Tannier, 2018). We can draw a parallel with previous qualitative research on the possibilities of AI technologies within newsrooms where human journalists intend to keep the lead in these new forms of human-machine collaborations (Dierickx, 2020; Gutierrez Lopez et al., 2022).

Effective AFC tools require efficient technology. Nevertheless, no technology can be used sufficiently on its own, despite various technical challenges. Therefore, (re)connecting communities means responding to both technological and social issues that arise upstream and downstream of the automated process. Our observations provide useful reflections for further research in AFC, whether working on improving data quality, models for standardizing annotations, the explainability of processes and results, or improving human-computer interaction and the overall user experience to better support professional practices. Drawing links between the adjacent communities surrounding the BO also implies that is connects research fields because AFC involves different but complementary scientific disciplines.

References

Abdelnabi, S., Hasan, R., & Fritz, M. (2021). *Open-domain, content-based, multi-modal fact-checking of out-of-context images via online resources*. Retrieved from http://arxiv.org/abs/2112.00061

Adair, B., Li, C., Yang, J., & Yu, C. (2017, October 13–14). *Progress toward "the Holy Grail": The continued quest to automate fact-checking*. Paper presented at the Computation + Journalism Symposium, Evanston, IL. Retrieved from https://ranger.uta.edu/~cli/pubs/2017/factcheckingcj17-adair.pdf

Adler, B., & Boscaini-Gilroy, G. (2019). *Real-time claim detection from news articles and retrieval of semantically-similar factchecks*. Retrieved from http://arxiv.org/abs/1907.02030

Akkerman, S. F., & Bakker, A. (2011). Boundary crossing and boundary objects. *Review of Educational Research, 81*(2), 132–169. doi:10.3102/0034654311404435

Alhindi, T., Alabdulkarim, A., Alshehri, A., Abdul-Mageed, M., & Nakov, P. (2021). AraStance: A multicountry and multi-domain dataset of Arabic stance detection for fact-checking. In *Proceedings of the Fourth Workshop on NLP for Internet Freedom: Censorship, Disinformation, and Propaganda* (pp. 171–177). Stroudsburg, PA: Association for Computational Linguistics. doi:10.18653/v1/2021.nlp4if-1.9

Azevedo, L. (2018). Truth or lie: Automatically fact checking news. In *Companion Proceedings of the Web Conference 2018* (pp. 807–811). Geneva, Switzerland: International World Wide Web Conferences Steering Committee. doi:10.1145/3184558.3186567

Bañon Castellón, L. (2021). Audiovisual verification in the evolution of television newsrooms: Al Jazeera and the transition from satellite to the cloud. *Anàlisi, 64*(2021), 85–102. doi:10.5565/rev/analisi.3414

Bardoel, J., & d'Haenens, L. (2004). Media responsibility and accountability. New conceptualizations and practices. *Communications, 29*(1), 5–25. doi:10.1515/comm.2004.007

Bartneck, C., Lutge, C., Wagner, A., & Welsh, S. (2021). *An introduction to ethics in robotics and AI an introduction to ethics in robotics and AI*. Cham, Switzerland: Springer Nature.

Berendt, B., Burger, P., Hautekiet, R., Jagers, J., Pleijter, A., & Van Aelst, P. (2021). FactRank: Developing automated claim detection for Dutch-language fact-checkers. *Online Social Networks and Media, 22*(100113), 1–12. doi:10.1016/j.osnem.2020.100113

Boididou, C., Papadopoulos, S., Zampoglou, M., Apostolidis, L., Papadopoulou, O., & Kompatsiaris, Y. (2018). Detection and visualization of misleading content on Twitter. *International Journal of Multimedia Information Retrieval, 7*(1), 71–86. doi:10.1007/s13735-017-0143-x

Borges, L., Martins, B., & Calado, P. (2019). Combining similarity features and deep representation learning for stance detection in the context of checking fake news. *ACM Journal of Data and Information Quality, 11*(3), 1–26. doi:10.1145/3287763

Brand, E., Roitero, K., Soprano, M., & Demartini, G. (2021, October). E-BART: Jointly predicting and explaining truthfulness. In I. Augenstein, P. Papotti, & D. Wright (Eds.), *Proceedings of the 2021 Truth and Trust Online Conference* (pp. 18–27). Arlington, VA: Hacks Hackers. Retrieved from https://truthandtrustonline.com/wp-content/uploads/2021/10/TTO2021_paper_16-1.pdf

Çarik, B., & Yeniterzi, R. (2021). SU-NLP at CheckThat! 2021: Check-worthiness of Turkish tweets. In *Proceedings of CLEF 2021—Conference and Labs of the Evaluation Forum* (pp. 476–483). Bucharest, Romania: CEUR Workshop Proceedings. Retrieved from http://ceur-ws.org/Vol2936/paper-37.pdf

Cavaliere, P. (2020). From journalistic ethics to fact-checking practices: Defining the standards of content governance in the fight against disinformation. *Journal of Media Law, 12*(2), 133–165. doi:10.1080/17577632.2020.1869486

Cazalens, S., Lamarre, P., Leblay, J., Manolescu, I., & Tannier, X. (2018). A content management perspective on fact-checking. In *Proceedings of the Web Conference 2018* (pp. 565–574). New York, NY: ACM Press.

Çoğaltay, N., & Karadağ, E. (2015). Introduction to meta-analysis. In E. Karadağ (Ed.), *Leadership and organizational outcomes: Meta-analysis of empirical studies* (pp. 19–28). Cham, Switzerland: Springer International Publishing.

Davis, J., Mengersen, K., Bennett, S., & Mazerolle, L. (2014). Viewing systematic reviews and metaanalysis in social research through different lenses. *SpringerPlus, 3*(1), 1–9. doi:10.1186/21931801-3-511 de Haan, Y., Van Den Berg, E., Goutier, N., Kruikemeier, S., & Lecheler, S. (2022). Invisible friend or foe? How journalists use and perceive algorithmic-driven tools in their research process. *Digital Journalism, 10*(10), 1775–1793. doi:10.1080/21670811.2022.2027798

Demartini, G., Mizzaro, S., & Spina, D. (2020). Human-in-the-loop artificial intelligence for fighting online misinformation: Challenges and opportunities. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, 43*(3), 65–74. Retrieved from http://sites.computer.org/debull/A20sept/p65.pdf

Denaux, R., & Gomez-Perez, J. M. (2020). Linked credibility reviews for explainable misinformation detection. In J. Z. Pan, V. Tamma, C. D'Amato, K. Janowicz, B. Fu, A. Polleres, . . . L. Kagal (Eds.), *Lecture notes in computer science* (pp. 147–163). Cham, Switzerland: Springer International Publishing.

Dhankar, A., Zaïane, O. R., & Bolduc, F. (2022). UofA-Truth at Factify 2022: Transformer and transfer learning based multi-modal fact-checking. *ArXiv*. doi:10.48550/arXiv.2203.07990

Diakopoulos, N. (2019). *Automating the news: How algorithms are rewriting the media*. Cambridge, MA: Harvard University Press.

Diakopoulos, N. (2020). Computational news discovery: Towards design considerations for editorial orientation algorithms in journalism. *Digital Journalism, 8*(7), 945–967. doi:10.1080/21670811.2020.1736946

Diakopoulos, N., Trielli, D., & Lee, G. (2021). Towards understanding and supporting journalistic practices using semi-automated news discovery tools. *Proceedings of the ACM on Human-Computer Interaction, 5(CSCW2)*, 1–30. doi:10.1145/3479550

Dierickx, L. (2020). The social construction of news automation and the user experience. *Brazilian Journalism Research, 16*(3), 432–457. doi:10.25200/BJR.v16n3.2021.1289

Dietterich, T. (1995). Overfitting and undercomputing in machine learning. *ACM Computing Surveys, 27*(3), 326–327. doi:10.1145/212094.212114

Du, Y., Bosselut, A., & Manning, C. D. (2022). Synthetic disinformation attacks on automated fact verification systems. *Proceedings of the AAAI Conference on Artificial Intelligence, 36*(10), 10581–10589. doi:10.1609/aaai.v36i10.21302

Durier da Silva, F. C., Bicharra Garcia, A. C., & Matsui Siqueira, S. W. (2022). A systematic literature mapping on profile trustworthiness in fake news spread. In *2022 IEEE 25th International Conference on Computer*

Supported Cooperative Work in Design (CSCWD) (pp. 275–279). Hangzhou, China: IEEE. doi:10.1109/CSCWD54268.2022.9776232

Ebadi, N., Jozani, M., Choo, K.-K. R., & Rad, P. (2022). A memory network information retrieval model for identification of news misinformation. *IEEE Transactions on Big Data, 8*(5), 1358–1370. doi:10.1109/TBDATA.2020.3048961

Fairbanks, J. P., Fitch, N., Bradfield, F., & Briscoe, E. (2020). Credibility development with knowledge graphs. In C. Grimme, M. Preuss, F. W. Takes, & A. Waldherr (Eds.), *Disinformation in open online media* (pp. 33–47). Cham, Switzerland: Springer International Publishing. doi:10.1007/978-3-030-39627-5_4

Frost, C. (2015). *Journalism ethics and regulation*. London, UK: Routledge. doi:10.4324/9781315757810

Fox, N. J. (2011). Boundary objects, social meanings and the success of new technologies. *Sociology, 45*(1), 70–85. doi:10.1177/0038038510387196

Gad-Elrab, M. H., Stepanova, D., Urbani, J., & Weikum, G. (2019). ExFaKT: A framework for explaining facts over knowledge graphs and text. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining* (pp. 87–95). New York, NY: Association for Computing Machinery. doi:10.1145/3289600.3290996

Gallofré Ocaña, M., & Opdahl, A. L. (2020). Challenges and opportunities for journalistic knowledge platforms. *Conference on Information and Knowledge Management (CIKM 2020 Workshops)*. Retrieved from https://ceur-ws.org/Vol-2699/paper43.pdf

Gao, J., Hoffmann, H.-F., Oikonomou, S., Kiskovski, D., & Bandhakavi, A. (2022). *Logically at Factify 2022: Multimodal fact verification*. Retrieved from http://arxiv.org/abs/2112.09253

Gencheva, P., Koychev, I., Màrquez, L., Barrón-Cedeño, A., & Nakov, P. (2019). *A context-aware approach for detecting check-worthy claims in political debates*. Retrieved from http://arxiv.org/abs/1912.08084

Graves, L. (2018). Boundaries not drawn: Mapping the institutional roots of the global fact-checking movement. *Journalism Studies, 19*(5), 613–631. doi:10.1080/1461670x.2016.1196602

Graves, L., & Cherubini, F. (2016). *The rise of fact-checking sites in Europe*. Oxford, UK: Reuters Institute for the Study of Journalism.

Grimmer, J., Roberts, M. E., & Stewart, B. M. (2021). Machine learning for social science: An agnostic approach. *Annual Review of Political Science, 24*, 395–419. doi:10.1146/annurev-polisci-053119015921

Guess, A. M., Lockett, D., Lyons, B., Montgomery, J. M., Nyhan, B., & Reifler, J. (2020). "Fake news" may have limited effects beyond increasing beliefs in false claims. *Harvard Kennedy School Misinformation Review, 1*(1). doi:10.37016/mr-2020-004

Guo, Z., Schlichtkrull, M., Thorne, J., Vlachos, A., Christodoulopoulos, C., Cocarascu, O., & Mittal, A. (2022). A survey on automated fact-checking. *Transactions of the Association for Computational Linguistics, 10*(2022), 178–206. doi:10.1162/tacl_a_00454

Gupta, N., Patel, H., Afzal, S., Panwar, N., Mittal, R. S., Guttula, S., . . . Saha, D. (2021). *Data Quality Toolkit: Automatic assessment of data quality and remediation for machine learning datasets*. Retrieved from http://arxiv.org/abs/2108.05935

Gutierrez Lopez, M., Porlezza, C., Cooper, G., Makri, S., MacFarlane, A., & Missaoui, S. (2022). A question of design: Strategies for embedding AI-driven tools into journalistic work routines. *Digital Journalism, 11*(3), 484–503. doi:10.1080/21670811.2022.2043759

Hanselowski, A., & Gurevych, I. (2017, November). A framework for automated fact-checking for realtime validation of emerging claims on the web. In *NIPS 2017 Workshop on Prioritising Online Content* (pp. 1–3). Long Beach, CA: NIPS. Retrieved from https://www.k4all.org/wpcontent/uploads/2017/09/WPOC2017_paper_6.pdf

Hassan, N., Zhang, G., Arslan, F., Caraballo, J., Jimenez, D., Gawsane, S., … Tremayne, M. (2017). ClaimBuster: The first-ever end-to-end fact-checking system. *Proceedings of the VLDB Endowment International Conference on Very Large Data Bases, 10*(12), 1945–1948. doi:10.14778/3137765.3137815

Hoque, M. A., Ferdous, M. S., Khan, M., & Tarkoma, S. (2021). Real, forged or deep fake? Enabling the ground truth on the internet. *IEEE Access: Practical Innovations, Open Solutions, 9*(2021), 160471–160484. doi:10.1109/access.2021.31315170

Horne, B. D., Dron, W., Khedr, S., & Adali, S. (2018). Assessing the news landscape: A multi-module toolkit for evaluating the credibility of news. In *Companion Proceedings of the Web Conference 2018* (pp. 235–238). Lyon, France: International World Wide Web Conferences Steering Committee. doi:10.1145/3184558.3186987

Hui Xian Ng, L., & Carley, K. M. (2021). "The coronavirus is a bioweapon": Classifying coronavirus stories on fact-checking sites. *Computational and Mathematical Organization Theory, 27*(2), 179–194. doi:10.1007/s10588-021-09329-w

Huvila, I., Anderson, T. D., Jansen, E. H., Mckenzie, P. J., & Worrall, A. (2017). Boundary objects in information science. *Journal of the Association for Information Science and Technology, 68*(8), 1807–1822. doi:10.1002/asi.23817

Jaroucheh, Z., Alissa, M., Buchanan, W. J., & Liu, X. (2020). TRUSTD: Combat fake content using blockchain and collective signature technologies. In *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)* (pp. 1235–1240). Madrid, Spain: IEEE. doi:10.1109/COMPSAC48688.2020.00-87

Jiang, S., Baumgartner, S., Ittycheriah, A., & Yu, C. (2020). Factoring fact-checks: Structured information extraction from fact-checking articles. In *Proceedings of the Web Conference 2020* (pp. 1592– 1603). New York, NY: ACM. doi:10.1145/3366423.3380231

Kalsnes, B. (2018). Fake news. In *Oxford Research Encyclopedia of Communication*. Oxford, UK: Oxford University Press. doi:10.1093/acrefore/9780190228613.013.809

Kar, D. (2020). Spotting misinformation to limit the impact of disruption on society by using machine learning. In *2020 IEEE Applied Signal Processing Conference (ASPCON)* (pp. 26–30). Kolkata, India: APSCON. doi:10.1109/ASPCON49795.2020.9276723

Katsaounidou, A. N., Gardikiotis, A., Tsipas, N., & Dimoulas, C. A. (2020). News authentication and tampered images: Evaluating the photo-truth impact through image verification algorithms. *Heliyon, 6*(12), 1–22. doi:10.1016/j.heliyon.2020.e05808

Kläs, M., & Vollmer, A. M. (2018). Uncertainty in machine learning applications: A practice-driven classification of uncertainty. In *Developments in language theory* (pp. 431–438). Cham, Switzerland: Springer International Publishing.

Komendantova, N., Ekenberg, L., Svahn, M., Larsson, A., Shah, S. I. H., Glinos, M., . . . Danielson, M. (2021). A value-driven approach to addressing misinformation in social media. *Humanities and Social Sciences Communications, 8*(1), 1–12. doi:10.1057/s41599-020-00702-9

Konstantinovskiy, L., Price, O., Babakar, M., & Zubiaga, A. (2021). Toward automated factchecking: Developing an annotation schema and benchmark for consistent automated claim detection. *Digital Threats: Research and Practice, 2*(2), 1–16. doi:10.1145/3412869

Kotonya, N., & Toni, F. (2020). Explainable automated fact-checking for public health claims. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 7740–7754). Stroudsburg, PA: Association for Computational Linguistics. doi:10.18653/v1/2020.emnlp-main.623

Lakshmanan, L. V. S., Simpson, M., & Thirumuruganathan, S. (2019). Combating fake news: A data management and mining perspective. *Proceedings of the VLDB Endowment, 12*(2), 1990–1993. doi:10.14778/3352063.3352117

Lazarski, E., Al-Khassaweneh, M., & Scotts Howard, C. (2021). Using NLP for fact checking: A survey. *Designs, 5*(3), 1–22. doi:10.3390/designs5030042

Lease, M. (2011). On quality control and machine learning in crowdsourcing. In *Workshops at the TwentyFifth AAAI Conference on Artificial Intelligence* (pp. 97–102). San Francisco, CA: AAAI.

Lewis, S. C., & Usher, N. (2016). Trading zones, boundary objects, and the pursuit of news innovation: A case study of journalists and programmers. *Convergence, 22*(5), 543–560. doi:10.1177/1354856515623865

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., . . . Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Annals of Internal Medicine, 151*(4), 65–94. doi:10.7326/0003-4819-151-4-200908180-00136

Majithia, S., Arslan, F., Lubal, S., Jimenez, D., Arora, P., Caraballo, J., & Li, C. (2019). ClaimPortal: Integrated monitoring, searching, checking, and analytics of factual claims on twitter. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (pp. 153–158). Florence, Italy: Association for Computational Linguistics. doi:10.18653/v1/p19-3026

Masood, R., & Aker, A. (2018). The fake news challenge: Stance detection using traditional machine learning approaches. In *Proceedings of the 10th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management* (pp. 128–135), Seville, Spain: KMIS. doi:10.5220/0006898801280135

Mena, P. (2019). Principles and boundaries of fact-checking: Journalists' perceptions. *Journalism Practice, 13*(6), 657–672. doi:10.1080/17512786.2018.1547655

Mengist, W., Soromessa, T., & Legese, G. (2020). Method for conducting systematic literature review and meta-analysis for environmental science research. *MethodsX, 7*(100777), 1–11. doi:10.1016/j.mex.2019.100777

Micallef, N., Armacost, V., Memon, N., & Patil, S. (2022). True or false: Studying the work practices of professional fact-checkers. *Proceedings of the ACM on Human–Computer Interaction, 6*(CSCW1), 1–44. doi:10.1145/3512974

Miceli, M., Posada, J., & Yang, T. (2021). *Studying up machine learning data: Why talk about bias when we mean power?* Retrieved from http://arxiv.org/abs/2109.08131

Middleton, S. E., Papadopoulos, S., & Kompatsiaris, Y. (2018). Social computing for verifying social media content in breaking news. *IEEE Internet Computing, 22*(2), 83–89. doi:10.1109/mic.2018.112102235

Miranda, S., Nogueira, D., Mendes, A., Vlachos, A., Secker, A., Garrett, R., . . . Marinho, Z. (2019). Automated fact checking in the newsroom. In *The World Wide Web Conference* (pp. 3579–3583). New York, NY: ACM. doi:10.1145/3308558.3314135

Molina, M., & Garip, F. (2019). Machine learning for sociology. *Annual Review of Sociology, 45*(1), 27–45. doi:10.1146/annurev-soc-073117-041106

Nakov, P., Corney, D. P. A., Hasanain, M., Alam, F., Elsayed, T., Barrón-Cedeño, A., . . . Da San Martino, G. (2021). Automated fact-checking for assisting human fact-checkers. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence* (pp. 4826–4832). Montreal, Canada: IJCAI. Retrieved from https://arxiv.org/abs/2103.07769

Nguyen, A. T., Kharosekar, A., Krishnan, S., Krishnan, S., Tate, E., Wallace, B. C., & Lease, M. (2018). Believe it or not: Designing a human-AI partnership for mixed-initiative fact-checking. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (pp. 189–199). New York, NY: Association for Computing Machinery. doi:10.1145/3242587.3242666

Northcutt, C. G., Athalye, A., & Mueller, J. (2021). *Pervasive label errors in test sets destabilize machine learning benchmarks*. arXiv Preprint. doi:10.48550/arXiv.2103.14749

Nygren, T., Guath, M., Axelsson, C.-A. W., & Frau-Meigs, D. (2021). Combatting visual fake news with a professional fact-checking tool in education in France, Romania, Spain, and Sweden. *Information, 12*(5), 1–25. doi:10.3390/info12050201

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery, 88*(105906), 1–9. doi:10.1016/j.ijsu.2021.105906

Papadopoulou, O., Zampoglou, M., Papadopoulos, S., & Kompatsiaris, I. (2019). A corpus of debunked and verified user-generated videos. *Online Information Review, 43*(1), 72–88. doi:10.1108/oir03-2018-0101

Pathak, A., & Srihari, R. (2019). BREAKING! Presenting fake news corpus for automated fact checking. In *Proceedings of the 57th annual meeting of the association for computational linguistics: Student research workshop* (pp. 357–362). Florence, Italy: Association for Computational Linguistics. doi:10.18653/v1/p19-2050

Petticrew, M., & Roberts, H. (2006). *Systematic Reviews in the social sciences: A practical guide*. Malden, MA: Wiley-Blackwell.

Picha Edwardsson, M., Al-Saqaf, W., & Nygren, G. (2021). Verification of digital sources in Swedish newsrooms—A technical issue or a question of newsroom culture. *Journalism Practice*, 1–18. doi:10.1080/17512786.2021.2004200

Posetti, J., & Matthews, A. (2018). *A short guide to the history of "fake news" and disinformation*. Washington, DC: International Center for Journalists. Retrieved from https://www.icfj.org/sites/default/files/201807/A%20Short%20Guide%20to%20History%20of%20Fake%20News%20and%20Disinformation_IC FJ%20Final.pdf

Qaddoura, R., Faris, H., & Aljarah, I. (2020). An efficient clustering algorithm based on the K-Nearest neighbors with an indexing ratio. *International Journal of Machine Learning and Cybernetics, 11*(3), 675–714. doi:10.1007/s13042-019-01027-z

Ramage, D., Rosen, E., Chuang, J., Manning, C. D., & Mcfarland, D. A. (2009). Topic modeling for the social sciences. *NIPS 2009 Workshop on Applications for Topic Models: Text and Beyond, 5*(27), 1–4. Retrieved from https://nlp.stanford.edu/pubs/tmt-nips09.pdf

Reis, J. C. S., & Benevenuto, F. (2021). Supervised learning for misinformation detection in WhatsApp. In *Proceedings of the Brazilian Symposium on Multimedia and the Web* (pp. 245–252). New York, NY: Association for Computing Machinery. doi:10.1145/3470482.3479641

Reis, J. C. S., Correia, A., Murai, F., Veloso, A., & Benevenuto, F. (2019). Supervised learning for fake news detection. *IEEE Intelligent Systems, 34*(2), 76–81. doi:10.1109/mis.2019.2899143

Rony, M. M. U., Hoque, E., & Hassan, N. (2020). ClaimViz: Visual analytics for identifying and verifying factual claims. In *2020 IEEE Visualization Conference (VIS)* (pp. 246–250). New York, NY: IEEE. doi:10.1109/VIS47514.2020.00056

Saeed, M., & Papotti, P. (2021). Fact-checking statistical claims with Tables. *IEEE Data Engineering Bulletin, 44*(3), 27–38. Retrieved from http://sites.computer.org/debull/A21sept/p27.pdf

Sepúlveda-Torres, R., Bonet-Jover, A., & Saquete, E. (2021). "Here are the rules: Ignore all rules": Automatic contradiction detection in Spanish. *Applied Sciences, 11*(7), 1–15. doi:10.3390/app11073060

Shelby, L. B., & Vaske, J. J. (2008). Understanding meta-analysis: A review of the methodological literature. *Leisure Sciences, 30*(2), 96–110. doi:10.1080/01490400701881366

Shi, L., Bhattacharya, N., Das, A., Lease, M., & Gwizdka, J. (2022). The effects of interactive AI design on user behavior: An eye-tracking study of fact-checking COVID-19 claims. In *ACM SIGIR Conference on Human Information Interaction and Retrieval* (pp. 315–320). New York, NY: ACM. doi:10.1145/3498366.3505786

Silva, R. M., Santos, R. L. S., Almeida, T. A., & Pardo, T. A. S. (2020). Towards automatically filtering fake news in Portuguese. *Expert Systems with Applications, 146*(113199), 1–14. doi:10.1016/j.eswa.2020.113199

Singer, J. B. (2018). Fact-checkers as entrepreneurs: Scalability and sustainability for a new form of watchdog journalism. *Journalism Practice, 12*(8), 1070–1080. doi:10.1080/17512786.2018.1493946

Singer, J. B. (2021). Border patrol: The rise and role of fact-checkers and their challenge to journalists' normative boundaries. *Journalism, 22*(8), 1929–1946. doi:10.1177/1464884920933137

Singh, P., Das, A., Li, J. J., & Lease, M. (2021). *The case for claim difficulty assessment in automatic fact checking*. Retrieved from http://arxiv.org/abs/2109.09689

Sirén-Heikel, S., Kjellman, M., & Lindén, C.-G. (2022). At the crossroads of logics: Automating newswork with artificial intelligence—(Re)defining journalistic logics from the perspective of technologists. *Journal of the Association for Information Science and Technology, 74*(3), 354–366. doi:10.1002/asi.24656

Školkay, A., & Filin, J. (2019). A comparison of fake news detecting and fact-checking AI based solutions. *Studia Medioznawcze, 20*(4), 365–383. doi:10.33077/uw.24511617.ms.2019.4.187

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*(2019), 333–339. doi:10.1016/j.jbusres.2019.07.039

Stray, J. (2019). Making artificial intelligence work for investigative journalism. *Digital Journalism, 7*(8), 1076–1097. doi:10.1080/21670811.2019.1630289

Svahn, M., & Perfumi, S. C. (2021). A conceptual model for approaching the design of anti-disinformation tools. In *International Conference on Electronic Participation* (pp. 66–76). Cham, Switzerland: Springer. doi:10.1007/978-3-030-82824-0_6

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology, 8*(45), 1–10. doi:10.1186/1471-2288-8-45

Thorne, J., Vlachos, A., Christodoulopoulos, C., & Mittal, A. (2018). FEVER: A large-scale dataset for fact extraction and verification. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (Vol 1: Long Papers, pp. 809–819). Stroudsburg, PA: Association for Computational Linguistics. doi:10.18653/v1/n18-1074

Trompette, P., & Vinck, D. (2009). Revisiting the notion of boundary object. *Revue d'Anthropologie des Connaissances, 3*(1), a–v. doi:10.3917/rac.006.0003

Umarova, K., & Mustafaraj, E. (2019). How partisanship and perceived political bias affect Wikipedia entries of news sources. In *Companion Proceedings of the 2019 World Wide Web Conference* (pp. 1248–1253). New York, NY: ACM. doi:10.1145/3308560.3316760

Vayansky, I., & Kumar, S. A. P. (2020). A review of topic modeling methods. *Information Systems, 94*(2020). doi:10.1016/j.is.2020.101582

Vizoso, Á., Vaz-Álvarez, M., & López-García, X. (2021). Fighting deepfakes: Media and internet giants' converging and diverging strategies against hi-tech misinformation. *Media and Communication, 9*(1), 291–300. doi:10.17645/mac.v9i1.3494

Vlachos, A., & Riedel, S. (2014). Fact checking: Task definition and dataset construction. In *Proceedings of the ACL 2014 Workshop on Language Technologies and Computational Social Science* (pp. 18–22). Stroudsburg, PA: Association for Computational Linguistics. doi:10.3115/v1/w14–2508

Vryzas, N., Katsaounidou, A., Vrysis, L., Kotsakis, R., & Dimoulas, C. (2022). A prototype web application to support human-centered audiovisual content authentication and crowdsourcing. *Future Internet, 14*(3), 1–17. doi:10.3390/fi14030075

Wang, W. Y. (2017). "Liar, liar pants on fire": A new benchmark dataset for fake news detection. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics* (Vol. 2: Short Papers, pp. 422–426). Vancouver, Canada: Association for Computational Linguistics. doi:10.18653/v1/p17-2067

Welbers, K., Van Atteveldt, W., & Benoit, K. (2017). Text analysis in R. *Communication Methods and Measures, 11*(4), 245–265. doi:10.1080/19312458.2017.1387238

Yang, K.-C., Varol, O., Davis, C. A., Ferrara, E., Flammini, A., & Menczer, F. (2019). Arming the public with artificial intelligence to counter social bots. *Human Behavior and Emerging Technologies, 1*(1), 48–61. doi:10.1002/hbe2.115

Yang, Y., Zheng, L., Zhang, J., Cui, Q., Li, Z., & Yu, P. S. (2018). *TI–CNN: Convolutional neural networks for fake news detection*. Retrieved from http://arxiv.org/abs/1806.00749

Zeng, X., Abumansour, A. S., & Zubiaga, A. (2021). Automated fact checking: A survey. *Language and Linguistics Compass, 15*(10), 1–21. doi:10.1111/lnc3.12438

Zhou, J., Hu, H., Li, Z., Yu, K., & Chen, F. (2019). Physiological indicators for user trust in machine learning with influence enhanced fact-checking. In A. Holzinger, P. Kieseberg, A. Tjoa, & E. Weippl (Eds.), *Machine learning and knowledge extraction* (pp. 73–88). Cham, Switzerland: Springer. doi:10.1007/978-3-030-29726-8_7

Zlatkova, D., Nakov, P., & Koychev, I. (2019). Fact-checking meets fauxtography: Verifying claims about images. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing* (pp. 2099–2108).

Stroudsburg, PA: Association for Computational Linguistics. doi:10.18653/v1/d19-1216