

aedFaCT: Scientific Fact-Checking Made Easier via Semi-Automatic Discovery of Relevant Expert Opinions

Enes Altuncu^{1*}, Jason R.C. Nurse¹, Meryem Bagriacik², Sophie Kaleba², Haiyue Yuan¹, Lisa Bonheme², Shujun Li^{1*}

¹Institute of Cyber Security for Society (ICSS) & School of Computing, University of Kent

²School of Computing, University of Kent

Abstract

In this highly digitised world, fake news is a challenging problem that can cause serious harm to society. Considering how fast fake news can spread, automated methods, tools and services for assisting users to do fact-checking (i.e., fake news detection) become necessary and helpful, for both professionals, such as journalists and researchers, and the general public such as news readers. Experts, especially researchers, play an essential role in informing people about truth and facts, which makes them a good proxy for non-experts to detect fake news by checking relevant expert opinions and comments. Therefore, in this paper, we present *aedFaCT*, a web browser extension that can help professionals and news readers perform fact-checking via the automatic discovery of expert opinions relevant to the news of concern via shared keywords. Our initial evaluation with three independent testers (who did not participate in the development of the extension) indicated that *aedFaCT* can provide a faster experience to its users compared with traditional fact-checking practices based on manual online searches, without degrading the quality of retrieved evidence for fact-checking. The source code of *aedFaCT* is publicly available at <https://github.com/altuncu/aedFaCT>.

1 Introduction

The digital age has evolved into an infodemic age, with the rapid propagation of false and misleading information in this highly digitised world, mixed with true and reliable information. As part of this, fake news prevents society from obtaining accurate information based on real evidence. The COVID-19 pandemic has demonstrated how fake news can seriously cause harm to people (BBC News 2020).

Considering the amount of information available online and how fast information can be widely disseminated with the help of digital communication technologies, detecting fake news at scale is an important task. Therefore, many researchers have studied automated fact-checking methods (Zhou and Zafarani 2020). However, automated fact-checking solutions are yet to be sufficient to adapt to various contexts, languages, and modalities. In addition, they insufficiently consider human factors, such as trust and usability,

which is crucial for practical use (Das et al. 2023). These paved the way for semi-automated solutions that attempt to combine human and machine intelligence. With this respect, the literature involves many fact-checking systems leveraging human-machine teaming in different ways (Guo et al. 2020; Das et al. 2023). Besides, there exists a wide range of tools and services that can assist professionals and common readers with fake news detection (Nakov et al. 2021).

Experts play a crucial role in the fight against fake scientific news by enlightening society with truths and facts through various communication channels, especially the news media. In the complicated landscape of the infodemic age, people are likely to seek help from experts they trust, such as scientists and professionals, since they are often considered the highly trusted groups in society (Ipsos MORI 2022). This makes them a proxy for non-experts to fact-check suspicious scientific claims. Other than giving expert comments and being interviewed, experts support journalists in reporting online false information to bridge gaps in their contextual understanding and methodological expertise (McClure Haughey, Povo, and Starbird 2022). Besides, experts are a crucial element of the fact-checking process conducted by human fact-checkers (Graves 2017). With this respect, there has been much effort to engage with experts for scientific fact-checking. To exemplify, Science Media Centre¹ aims to build bridges between experts and journalists so that scientific information covered in the media becomes accurate and evidence-based. Another example is Meedan’s Digital Health Lab², which is composed of scientists, content moderation experts, and journalists to support evidence-based responses to health misinformation.

However, this expert-journalist collaboration could be insufficient to combat fake scientific news due to several reasons, including challenges in scientific communication (Bucchi 2017), the existence of outlier experts who do not share the majority opinions, and the selection of experts with incompatible expertise (Palmer 2020). These problems indicate the need and potential usefulness of tools that can leverage multiple experts’ opinions as evidence for fact-checking purposes. Therefore, in this work, we present *aedFaCT*, a web browser extension that can help professionals

*Corresponding co-authors: Enes Altuncu (ea483@kent.ac.uk) and Shujun Li (S.J.Li@kent.ac.uk).

Copyright © 2023, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

¹<https://www.sciencemediacentre.org/>

²<https://meedan.com/programs/digital-health-lab>

and common readers to discover the opinions of multiple experts on relevant topics of a particular scientific news article in a semi-automated manner. aedFaCT extracts expert opinions from several credible news sources based on a number of candidate keywords automatically extracted from the target news article, and it also automatically retrieves relevant peer-reviewed scientific publications based on such keywords. Based on the results, users can make a decision on the veracity of suspicious claims on their own by considering the retrieved evidence. Moreover, aedFaCT enables users to see a list of researchers with relevant expertise based on their publications in order to inform them about who to follow and to approach on a specific topic. In a nutshell, aedFaCT is a “smart search assistant” for fact-checkers to help minimise the manual work they have to do using online search engines and other known information sources.

The rest of the paper is organised as follows. Section 2 briefly reviews related work. Then, Section 3 presents a focus group study to understand the mental process of users during scientific fact-checking. The architecture of the proposed system is introduced in Section 4, and the details of its evaluation are provided in Section 5. Finally, the paper is concluded with a brief discussion in Section 6 and the concluding remarks in Section 8.

2 Related Work

2.1 Human-Machine Teaming Approaches in Fact-Checking

Automated fact-checking at scale is a challenging task. Hence, recent research includes hybrid solutions based on human-machine teaming to assist fact-checkers and the general public with a level of automation in the process of fact-checking. For example, Nguyen et al. (2018) designed a mixed-initiative approach to fact-checking where the system predicts the veracity of a claim based on relevant articles with their stance towards the veracity of the claim and the reputation of each source. The users’ role in this design is to change the source reputation and stance of each article for more accurate prediction. More recently, Gupta et al. (2021) introduced an evidence retrieval approach to search for semantically-similar news articles to assist users when validating news articles. This system leaves the fact-checking decision to the user. Moreover, La Barbera, Roitero, and Mizzaro (2022) proposed a hybrid human-in-the-loop framework for the veracity assessment of claims, relying on three major components: AI, crowdsourcing, and experts. The veracity of the claim is considered as correctly classified if any component produces a prediction with a high confidence score. Otherwise, the claim is forwarded to the next component. Another human-in-the-loop AI system is HAMLET, a conceptual framework leveraging AI-expert teaming in multiple fact-checking tasks, such as the collection of expert data annotations and expert feedback, AI system performance monitoring, and life cycle management (Bandhakavi, Hoffmann, and Lear 2022). Finally, Arroyo Guardado et al. (2021) introduced a toolbox, namely Ms.W, combining several publicly available services and tools that help users with fact-checking and source credi-

bility assessment.

As another way of human-machine teaming, several fact-checking systems utilise crowd intelligence in different stages. For instance, Vo and Lee (2018) leveraged guardians, who are social media users correcting false information by referring to fact-checking URLs, and presented a fact-checking URL recommendation model to motivate them to engage more in fact-checking activities. Furthermore, social media companies enable users to flag posts containing false information and sent them to fact-checkers for further investigation if there are sufficient flags. Recently, Twitter launched Community Notes³ (previously known as Bird-watch), where users can add context to tweets to prevent the platform from false information.

2.2 Web Browser Extensions for Fact-Checking

Web browser extensions are quite useful for fact-checking, especially for web-based documents and articles. For example, *BRENDA* allows users to perform automatically fact-checking a news article or a snippet from the article (Botnevik, Sakariassen, and Setty 2020). It identifies the check-worthy claims, classifies them with a deep neural network, and then, shows the results to the user along with the evidence found from top-10 Google Search results. Another automated solution is *FADE*, which discovers multiple sources containing the same news story and performs automated fact-checking according to the trustworthiness of the news sources and the cited sources in the article (Jabiyev et al. 2021). Other than the solutions developed in academia, The Factual⁴ automatically rates news articles based on several characteristics, including their source quality and bias, author expertise, and tone.

There also exist Web browser extensions helping users with content analysis and evidence retrieval for fact-checking. One such tool is *InVID*, which helps users verify videos and images with a number of tools it contains (Teyssou 2019). As another example, *News2PubMed* retrieves relevant health research papers given a news article (Wang and Yu 2021). Another tool is called *News Scan*, which shows several characteristics of the source and content of news articles, such as source popularity, sentiment, objectivity, and bias, to assist users to make a judgement on the source and content credibility (Kevin et al. 2018). Finally, *NewsGuard*⁵ shows manually assigned source credibility ratings next to links on search engines and social media platforms.

3 Mental Process of Users During Fact-Checking

In this study, our aim is to develop a semi-automated fact-checking system for both professionals and common readers, which automates, at least, part of the users’ claim investigation process. To this end, we need to understand how users manually perform fact-checking and what strategies

³<https://help.twitter.com/en/using-twitter/community-notes>

⁴<https://www.thefactual.com/>

⁵<https://www.newsguardtech.com/>

they normally use to investigate a claim. From a general perspective, content is the most important factor for users during fact-checking (Pidikiti et al. 2020). Users mainly rely on their own knowledge and sense of judgement to make a decision, and they perform external acts of authentication (e.g., searching for more information via Google, family and friends, and experts) only if the first phase fails (Tandoc Jr. et al. 2018; Freiling 2019). When users seek external information, they commonly prefer information that they consider credible, such as peer-reviewed scientific papers, fact-checking reports, mainstream news articles, and Wikipedia entries (He and He 2022).

Since the current literature lacks a systematical discussion of how different processes that fact-checkers and common readers follow to verify scientific information, we conducted a focus group discussion between the first author and three other co-authors (the third, fourth and sixth) of this paper, who were all PhD students in Computer Science focusing on a relevant research topic (AI, NLP, and/or cyber security), to understand how users verify the veracity of news content. At the time of the discussion, only the first co-author knew about the details of the study as the initialiser of the work. During the discussion, an example news article containing a false claim about COVID-19 was provided to the participants, and the investigation of the claim has been performed by discussing each step of the fact-checking process. The discussion was conducted with three fact-checking scenarios, separately: (1) the participants (as researchers) performed fact-checking themselves; (2) the participants simulated how common readers with less domain knowledge would perform fact-checking without using expert opinions as a proxy; and (3) the participants simulated how common readers would perform fact-checking by using expert opinions as a proxy. For all the scenarios, the discussion was made with the same participants instead of separate groups of researcher and common reader participants, for the sake of simplicity and to allow cross-scenario alignment. Using researchers as common readers is not necessarily a problematic setup, since researchers are effectively like common readers for research areas beyond their own expertise (e.g., health and medicine for all the authors of this paper).

In the first scenario, the participants suggested identifying some keywords about the investigated claim and using them to search for relevant research papers on Google Scholar. Then, they suggested reading the abstracts of the first few publications to make a decision, provided that they trust the publisher. In the second scenario, however, they preferred to use Google Search to search for relevant material with the same set of keywords, assuming that common readers would have been unfamiliar with scientific papers and research databases. Then, they wanted to check out the search results that are trustworthy for them, e.g., a news article from a news outlet they trusted, or a post from a university’s official website advertising their research. Finally, in the third scenario, the participants suggested identifying multiple relevant domain experts through the websites of the corresponding institution or departments of well-known universities. Moreover, they found relevant news articles useful to identify some domain experts by checking who has been inter-

viewed in the article.

The focus group discussion provided three major conclusions on users’ scientific fact-checking process, supporting the findings of existing literature on the general fact-checking practices of fact-checkers and laypeople (Juneja and Mitra 2022; Micallef et al. 2022; He and He 2022): (1) domain experts were generally at the core of the fact-checking process, either explicitly, or implicitly through their publications; (2) only the sources they trusted were considered; and (3) multiple sources were taken into account for cross-checking what has been obtained.

4 System Design

4.1 Overview

The overview architecture of aedFaCT is shown in Figure 1. The system involves three main parts: (i) keyword extraction and selection; (ii) expert opinion discovery; (iii) scientific evidence retrieval.

4.2 Keyword Extraction and Selection

As the first step, the system needs to learn the context of the given news article by extracting a number of descriptive keywords. We designed this process as a human-in-the-loop mechanism to avoid topic drift while using the obtained keywords in searching. The system first fetches and parses the news content using the Newspaper3k⁶ library. Then, it performs automatic keyword extraction (AKE) with a state-of-the-art AKE algorithm, SIFRank+ (Sun et al. 2020), to obtain the initial set of keywords. Based on the findings of our previous study (Altuncu et al. 2022), we used our own version of SIFRank+, enhanced with post-processing. More precisely, the enhancement involves PoS-tagging-based filtering, and prioritising keywords contained in the corresponding domain thesaurus or Wikipedia as an entry. This ensures that only noun phrases are considered keywords, and contextual keywords are given priority. As AKE methods are incapable of providing sufficient accuracy (Papa-georgiou and Tsoumakas 2020), we ask users to select the keywords relevant to the article out of ten identified keywords through the pop-up window shown in the Web browser, as depicted in Figure 2. Users are also allowed to add and select their own keywords through the user interface.

4.3 Expert Opinion Discovery

This step aims to explore the scientific views or comments of domain experts in the news media on the identified topic. The system combines the keywords selected by the user in the previous step with the *AND* operator to generate a search query to search for relevant news items. Although a more useful query can be formed with a combination of different logical operators, we simply used the *AND* operator for the sake of simplicity. The searches are done via Google’s search APIs by considering the following types of news sources:

⁶<https://newspaper.readthedocs.io/en/latest/>

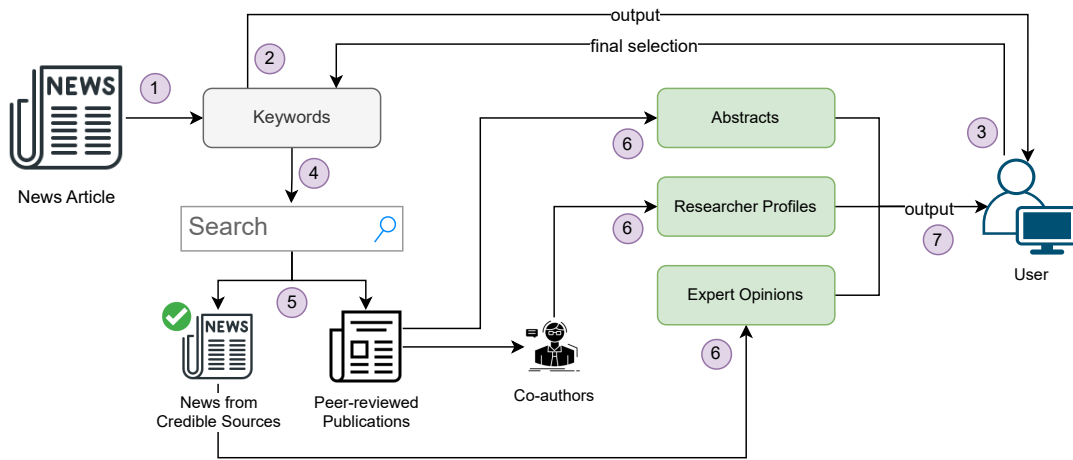


Figure 1: The architecture of aedFaCT

1. *Mainstream News Outlets*: This includes credible news outlets with high traffic and wide news coverage. We set up a Google site-restricted search engine, which allows 10 websites for inclusion, and covered 10 news outlets with high credibility in English-language and having no paywall. For the source credibility measure, we considered the Media Bias/Fact Check (MBFC) credibility ratings⁷ since it has been utilised by several recent studies (Krieg et al. 2020; Chen and Freire 2020; Weld, Glenski, and Althoff 2021). The included news outlets are shown in Table 1.
2. *Scientific News Outlets*: This type involves credible pro-science news websites, featuring scientific views and recent research findings. For this part, we set up another Google site-restricted search engine with 10 selected news websites. The selection was made according to the MBFC credibility ratings with the help of bias, credibility, and traffic filters, and the websites with pro-science bias, wide news coverage, higher traffic, and no paywall were preferred. Table 1 indicates the list of selected websites of this type.
3. *Other Credible News Sources*: In addition to the previous types, there are other types of news sources that might include expert opinions, such as news released by institutions and domain-specific news websites (e.g., Medscape, News Medical). To cover these, we set up a Google custom search engine without any site restriction to augment the search results containing the other two types of news sources. Since Google can also show results from non-news websites, we limited the search results with the *NewsArticle*⁸ Schema.org type to include only news articles. Furthermore, we utilised the Iffy Index of Unreliable Sources⁹, which is based on MBFC, to exclude untrustworthy news sources from the search results.

⁷<https://mediabiasfactcheck.com/>

⁸<https://schema.org/NewsArticle>

⁹<https://iffy.news/index/>

The search results obtained from the three search engines are aggregated with the given order. Although it is possible to merge the three search engines into a single custom search engine, we preferred to use site-restricted engines for the first two types of news sources since we observed that site-restricted search engines provide more reliable results, and custom search engines configured to search the entire Web are limited to a subset of the Google Web Search corpus¹⁰. Hence, we benefited from a custom search engine as a secondary source to populate the obtained results from the site-restricted search engines.

Once the aggregated set of search results is obtained, the system tries to capture expert opinions from each article, which are mostly in the form of reported speeches, since they contain the most indicative elements (e.g., reported speeches, named entities, and quotes) of page usefulness for fact-checking (Hasanain and Elsayed 2022). In this manner, it first downloads the news article with the Newspaper3k library. Then, the article is tokenised with two consecutive newline characters to obtain its paragraphs. Finally, for each paragraph, named entities are extracted with the spaCy library’s NER feature. Only the paragraphs which contain at least one person name, one academic organisation name (containing an indicative word or phrase, such as *university*, *institute*, *academy*, and *research centre*) and a pair of single or double quotation marks (indicating a reported speech) are selected. As an exception, the summary extracted by the Newspaper3k library is shown to users for the *The Conversation* news articles instead of retrieved expert opinions as they are already written by researchers and academics. As shown in Figure 3, the selected paragraphs are combined and shown to users in an individual box that also contains the source type (icon on the top-left), source name, and publish date. If the shown expert opinions are insufficient for a judgement and require further reading, users can click on the box to see the full article. Furthermore, a green clickable

¹⁰<https://support.google.com/programmable-search/answer/70392>

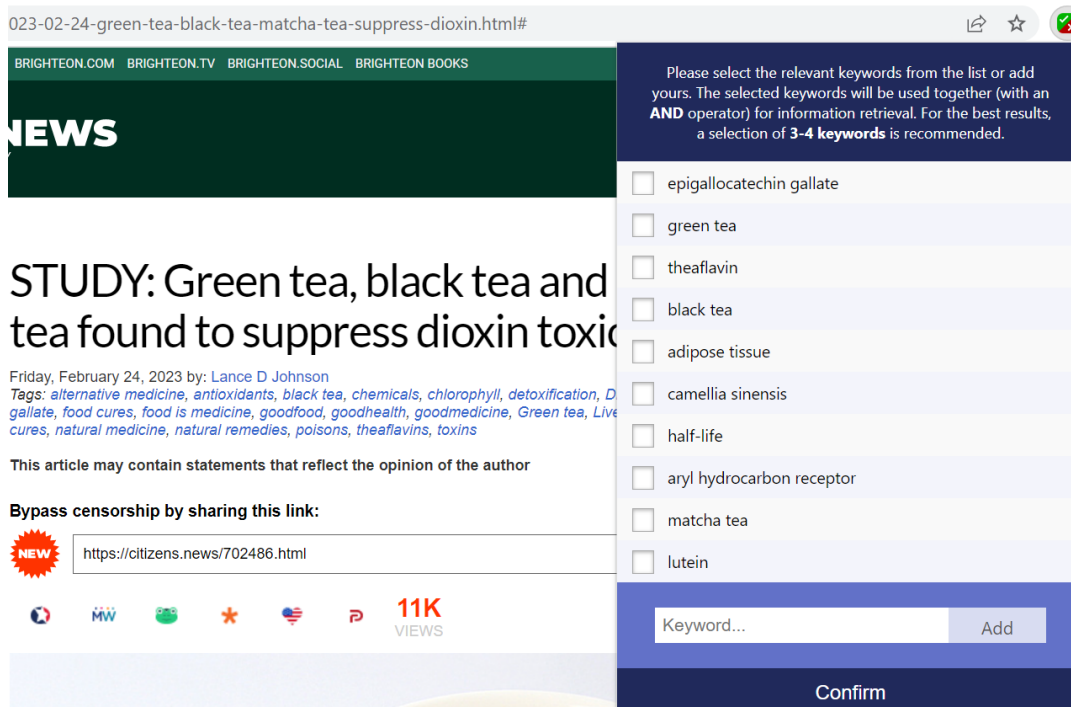


Figure 2: The user interface of the keyword extraction step in aedFaCT

Table 1: News outlets covered by the site-restricted search engines

Mainstream News Outlets	Scientific News Outlets
NPR (www.npr.org)	Science (www.science.org)
NBC News (www.nbcnews.com)	EurekAlert (www.eurekalert.org)
Sky News (news.sky.com)	The Scientist (www.the-scientist.com)
ABC News (www.abcnews.go.com)	Science News (www.sciencenews.org)
Euronews (www.euronews.com)	MIT Technology Review (www.technologyreview.com)
Reuters (www.reuters.com)	Popular Science (www.popsci.com)
BBC News (www.bbc.com)	Science Daily (www.sciencedaily.com)
PBS NewsHour (www.pbs.com/newshour)	Science Alert (www.sciencealert.com)
Associated Press (www.apnews.com)	Live Science (www.livescience.com)
CBS News (www.cbsnews.com)	The Conversation (www.theconversation.com)

tick directing to the corresponding MBFC credibility rating webpage is added next to the names of the mainstream and science news sources for better explainability.

4.4 Scientific Evidence Retrieval

Scientific publications can also be considered a source of expert opinions as they are written by domain experts. Therefore, this step aims to retrieve research papers relevant to the topic of the input article.

Similar to the previous step, we try to include only the records with high credibility. With this respect, we utilised Scopus API (with the help of the Pybliometrics library (Rose and Kitchin 2019)) to search for relevant peer-reviewed publications. The searches are made by combining the selected keywords with an *AND* operator, similar to the previous step. In addition, each keyword is surrounded by double quota-

tion marks since it enables the inclusion of loose matches by allowing for wildcards and lemmatisation (Beatty 2022). As shown in the upper side of Figure 4, the obtained search results are shown to the user inside individual boxes containing the title, source, publication year, and abstract, with an order of relevance and publication year.

In addition to the scientific evidence provided by the tool, users, especially fact-checkers and journalists, might want to know the experts on the topic themselves to follow their research and/or make contact with them. To enable this, our proposed tool profiles the co-authors of the publications retrieved in the previous step, by obtaining relevant information, e.g., profile links, from their Scopus and ORCID profiles. The obtained researcher profiles are ordered by their number of publications in the search result. In the case that this number is equal, they are ranked based on the amount

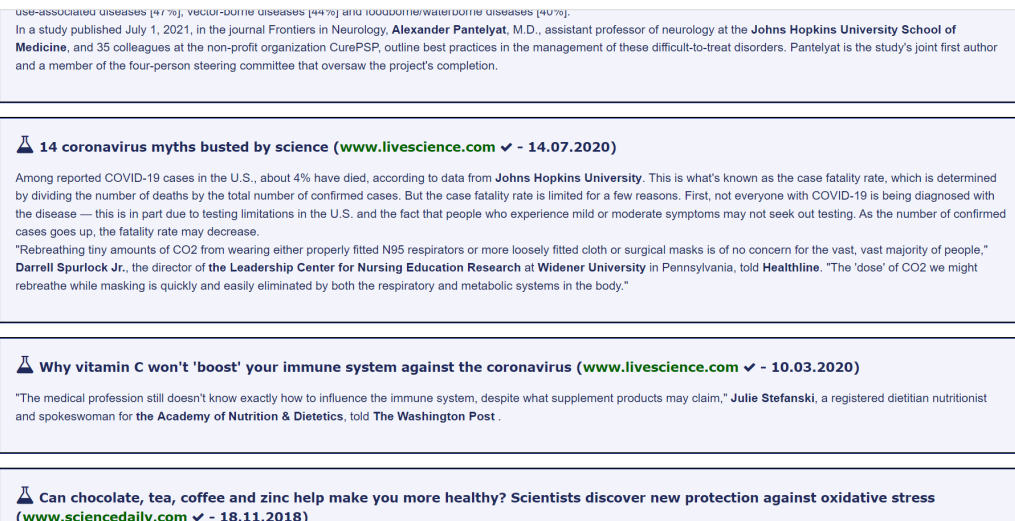


Figure 3: An example output from aedFaCT showing some of the retrieved news articles.

of information their profile contains to prioritise more contactable researchers. The bottom side of Figure 4 shows an example output from the user interface showing a list of researchers.

5 Evaluation

To check the functionality and validity of the proposed Web browser extension, we conducted an initial evaluation as a pilot study with three co-authors (the third, fourth, and fifth) of this paper (one male and two female researchers), who were not included in the design and implementation phases of the tool. They were provided with 20 health news articles released by multiple sources with different credibility levels. The health domain was selected in order for a better simulation of common readers since it was outside the participants' areas of expertise. Then, the participants were asked to investigate the veracity of each news article and provide ratings for the shown output, in two rounds: 1) manually by following the investigation practices in their daily lives, such as using a Web search engine, and/or using a research database; 2) by using our proposed tool, aedFaCT.

For collecting the ratings from the participants, we set up a survey on Google Forms, containing a rating scale for each processed news article in both rounds together with a figure explaining each option in the scale. In addition, the survey included two questions to assess the perceived success of aedFaCT in terms of which approach had been faster and more helpful (with the options *manual investigation*, *investigation with aedFaCT*, and *no difference*). Finally, it concluded with an open-ended question for comments and feedback.

In terms of the evaluation criteria in the rating scale, we followed Google's search quality guidelines (Google 2022), which was proposed for evaluating Google search engine results with human raters. Although there are criticisms regarding the inadequacy of such retrieval effectiveness tests (Lewandowski 2015), similar approaches are still

being used in the literature (Ciccone and Vickery 2015). The guidelines involve mainly two tasks: determining to what extent the page achieves its purpose ("*Page Quality*") and determining if search results are useful ("*Needs Met*"). Because our tool only benefits from credible news outlets and peer-reviewed publications, the former task is redundant in our case. Therefore, we only covered the latter task in our evaluation.

The "*Needs Met*" task involves two steps, which are about determining the user intent and the rating. Since all users of our tool will have the same intent, i.e., veracity assessment, the first step is redundant. Therefore, we only asked our evaluators to determine the rating of the search results by following the scale shown in Table 2.

As a result of the evaluation, the average rating of the three raters when they manually investigated the given news articles was 4.35. This average has risen to 4.57 when they utilised aedFaCT in their investigations. In addition, the raters were in moderate agreement that aedFaCT provided better or similar results with respect to what they were able to obtain with their manual investigations, with a Fleiss' Kappa of 53.33%. However, the raters have all agreed that fact-checking with aedFaCT was faster than their own practices. These results indicate that aedFaCT can help users perform fact-checking faster without degrading the quality of retrieved evidence for fact-checking. However, more extensive experiments are needed to evaluate its performance.

6 Further Discussions

6.1 Comparing aedFaCT with Existing Tools

aedFaCT differs from existing fact-checking systems in several ways. To begin with, to the best of our knowledge, it is the first fact-checking system completely based on expert opinion discovery although there exist studies leveraging experts in fact-checking (La Barbera, Roitero, and Mizzaro 2022; Bandhakavi, Hoffmann, and Lear 2022; Wang and Yu 2021). Secondly, it is an evidence retrieval tool, and the fi-

This is what I found...

News Articles	Scientific Publications	Researchers
<p>5 publications found</p>		
<p>A Detailed Comparison of the Use of Dietary Supplements before and during the COVID-19 Pandemic (Progress in Nutrition, 2022)</p> <p>Abstract: Mattioli 1885. Background and aim: Considering the coronavirus disease 2019 (COVID-19) pandemic and the expectation that dietary supplements (DS) boost the immune system of individuals, the aim of this research was to evaluate the use of DS and related factors by comparing periods before and during the pandemic. Methods: A descriptive cross-sectional internet-based study was conducted with 1488 participants from the general public aged above 19 years. Results: The median age of participants was 33.0 (19-69) years and 88% of participants were women. A total of 48.9% of participants (50.2% of females, 39.9% of males) used DS before COVID-19, and DS use during COVID-19 was reported to be 57.9% (58.6% of females, 52.8% males). Independent variables of DS use before and during COVID-19 were sex, age, body mass index (BMI), education, income, vitamin/mineral deficiency (VMD), alcohol use, and medication use. Participants who had no VMD used more DS, mostly herbal supplements such as grape seed, gotu kola, ginseng, ginkgo biloba and green tea. Vitamin D and C were the most commonly-used DS among participants with vitamin/mineral deficiency, followed by zinc and multivitamins. The main reason for DS use (47.2%) during COVID-19 was to strengthen immunity. Conclusions: The current findings may help to understand the preferences of individuals about DS use during pandemics. Understanding factors associated with the use of DS and their claimed immune-boosting effects may support future studies aiming to provide accurate information and motivate individuals towards healthy use of supplements during pandemics.</p> <p>Authors: Yonca Sevim (Department of Nutrition and Dietetics, Bahçeşehir Üniversitesi)</p>		
<p>Micronutrients and bioactive substances: Their potential roles in combating COVID-19 (Nutrition, 2021)</p> <p>Abstract: 2020 Elsevier Inc. Objectives: The coronavirus disease 2019 (COVID-19) pandemic is seriously threatening public health and setting off huge economic crises across the world. To the extent of possible, during COVID-19, there is a need to look for alternative strategies. Therefore, the aim of this research is to review the roles of micronutrients and bioactive substances in combating COVID-19.</p>		

This is what I found...

News Articles	Scientific Publications	Researchers
<p>11 researchers found</p>		
<p>Arman Arab (Isfahan University of Medical Sciences)</p> <p>https://www.researchgate.net/profile/Arman-Arab</p> <p>Fatemeh Shirani, Farzin Khorvash and Arman Arab, Review on selected potential nutritional intervention for treatment and prevention of viral infections: possibility of recommending these for Coronavirus 2019, International Journal of Food Properties, 23(1), pp. 1722-1736, (2020). doi:10.1080/10942912.2020.1825483.</p>		
<p>Yonca Sevim (Department of Nutrition and Dietetics, Bahçeşehir Üniversitesi)</p> <p>Yonca Sevim, A Detailed Comparison of the Use of Dietary Supplements before and during the COVID-19 Pandemic, Progress in Nutrition, 24(3), (no pages found), (2022). doi:10.23751/pn.v24i3.12493.</p>		
<p>Tibebe Selassie Seyoum Keflie (Institute of Biological Chemistry and Nutrition, Universität Hohenheim)</p>		

Figure 4: An example output from aedFaCT showing some of the retrieved scientific publications and their co-authors, respectively

Table 2: Rating scale for the Needs Met task (Google 2022)

Rating	Description
Fully Meets (<i>FullyM</i>)	All or almost all users would be immediately and fully satisfied by the result and would not need to view other results to satisfy their need.
Highly Meets (<i>HM</i>)	Very helpful for many or most users. Some users may wish to see additional results.
Moderately Meets (<i>MM</i>)	Helpful for many users OR very helpful for some users. Some or many users may wish to see additional results.
Slightly Meets (<i>SM</i>)	Helpful for fewer users. There is a connection between the query and the result, but not a strong or satisfying connection. Many or most users would wish to see additional results.
Fails to Meet (<i>FailsM</i>)	Completely fails to meet the needs of the users. All or almost all users would wish to see additional results.
Not Applicable (<i>N/A</i>)	The evaluator was unable to evaluate the result.

nal decision on the veracity is given by the user. Thus, it can establish trust among the users more easily unlike many fact-checking tools with a black-box design and fully automated decision-making mechanism, due to scepticism to-

wards automation (Juneja and Mitra 2022). Another strength of aedFaCT is that it targets both common readers and professionals by retrieving both news articles and scientific publications. This enables users to consider information sources

that they are more familiar with, depending on their level of expertise. In addition, it provides users with evidence from multiple sources and experts, which is a beneficial approach to breaking users out of their echo chambers. Finally, its overall workflow aligns with the common practices of human fact-checkers, in which engaging with experts is a key element, meaning that fact-checkers can use the tool to accelerate their claim investigation processes (Juneja and Mitra 2022; Micallef et al. 2022). To be more precise, an overview of different characteristics of aedFaCT and other existing tools is provided in Table 3. As a result, we believe that aedFaCT can be made a new useful tool for fighting against false information and has the potential to be a part of standard fact-checking processes performed by both human fact-checkers and common readers.

6.2 Limitations and Future Work

The existing version of the proposed tool has a number of limitations. Firstly, it depends on external APIs (i.e., Google and Scopus) having quotas for the number of requests. Google Custom Search API¹¹ allows 10,000 requests per day while Scopus Search APIs¹² have a weekly quota between 5,000 and 20,000 requests, depending on the used API service. This makes it quite difficult to deploy aedFaCT for a wider community. Another limitation of the tool is its relatively low speed during keyword extraction. Since AKE methods already suffer from poor accuracy (Papaianopoulou and Tsoumakas 2020), we preferred accuracy over speed when selecting the AKE method and used the one (i.e., SIFRank+) providing the best accuracy although there exist various lightweight AKE algorithms. Besides, the information retrieval process was based on the selected keywords combined simply with the *AND* operator. This causes fewer records as a result of the searches, especially when too many keywords were chosen by the user. Therefore, a smarter approach utilising a combination of logical operators (e.g., using the *OR* operator for similar keywords) is needed for obtaining better search results. Lastly, the evaluation of aedFaCT has been conducted as a pilot study with a small number of participants having similar backgrounds. Hence, more extensive experiments with a more diverse and representative participant population, covering both professionals (e.g., fact-checkers and journalists) and common readers, are required.

Apart from resolving the limitations, for future work, we aim to improve the capabilities of aedFaCT. The existing version does not specifically consider retrieving results from official websites, e.g., governmental organisations, NGOs, and academic institutions. These can be retrieved by checking the URL extensions of the general Web results. Moreover, we plan to incorporate research on claim detection into aedFaCT so that extracted keywords can be more focused on specific claims in the input article.

¹¹<https://developers.google.com/custom-search/v1/overview#pricing>

¹²https://dev.elsevier.com/api_key_settings.html

6.3 Broader Impact

This work has some potential outcomes from a broader perspective. Since it accelerates the fact-checking process for users, it might encourage them to make fact-checking a daily activity and increase awareness in society for tackling false information online. However, users should be conscious when assessing the veracity of news items with the expert opinions shown by aedFaCT. Although aedFaCT retrieves evidence only from trustworthy sources, the displayed expert opinions might contradict each other due to disagreements between different experts. Therefore, aedFaCT does not eliminate the need for critical thinking ability for its users.

7 Research Ethics Considerations

The work reported in this paper involved a focus group discussion participated and a validation experiment participated by some co-authors of the paper only. According to the research ethics guidelines of the University of Kent's Central Research Ethics Advisory Group and general advice given by the School of Computing's Research Ethics Officer, such user studies involving researchers who are part of the research only were exempted from going through a research ethics review process. Both user studies did not involve any explicit collection of personal data or other sensitive data, and all participants explicitly consented to participate. Participating in the studies did not cause any noticeable harm to participants, but brought some benefits to them – they could all achieve a better understanding of how to conduct fact-checking as a common reader and researcher.

8 Conclusion

Fake news is a challenging problem in society and causes serious harm. Its speed of propagation with the help of digital technologies suggests the need for automated solutions that can help people combat fake news at scale. Although there has been much effort to detect fake news, existing tools and services overlooked engaging with experts, who are commonly consulted during standard fact-checking processes. Therefore, this paper proposed *aedFaCT*, a Web browser extension that retrieves expert opinions related to a news article to help fact-checkers and the general public perform fact-checking. Our initial evaluation suggested that it can accelerate fact-checking process without negatively affecting the search quality.

CRedit authorship contribution statement **Enes Altuncu**: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing. **Jason Nurse**: Methodology, Writing – review & editing, Supervision. **Meryem Bagriacik**: Investigation, Validation, Writing – review & editing. **Sophie Kaleba**: Investigation, Validation, Writing – review & editing. **Haiyue Yuan**: Validation, Writing – review & editing. **Lisa Bonheme**: Investigation, Writing – review & editing. **Shujun Li**: Conceptualization, Methodology, Supervision, Writing – review & editing.

Acknowledgements We would like to thank all the reviewers for their valuable feedback. The first and third co-

Table 3: The comparison between aedFaCT and other existing web browser extensions for fact-checking

Tool	Approach	Task	Output	Domain
BRENDA	Automatic	Veracity Prediction	News articles	Any
FADE	Automatic	Veracity Prediction	News articles	Any
The Factual	Automatic	Credibility Assessment	Source & content credibility	Any
InVID	Semi-automatic	Content Analysis	Additional information about content	Any
News2PubMed	Automatic	Evidence Retrieval	Research papers	Health only
NewsGuard	Automatic	Credibility Assessment	Source credibility	Any
News Scan	Semi-automatic	Credibility Assessment	Source & content credibility	Any
aedFaCT	Semi-automatic	Evidence Retrieval	News articles, research papers & researchers	Any

authors, E. Altuncu and M. Bagriacik, were supported by funding from the Ministry of National Education, Republic of Turkey, through the MoNE-YLSY scholarship program.

References

- Altuncu, E.; Nurse, J. R. C.; Xu, Y.; Guo, J.; and Li, S. 2022. Improving Performance of Automatic Keyword Extraction (AKE) Methods Using PoS-Tagging and Enhanced Semantic-Awareness. arXiv:2211.05031 [cs.CL].
- Arroyo Guardño, D.; Gómez Espés, A.; Palmero Muñoz, S.; and Degli Esposti, S. 2021. On the design of a misinformation widget (Ms.W) against cloaked science. Technical report, Instituto de Física Aplicada (IFA), Ciencia y Tecnologías Físicas.
- Bandhakavi, A.; Hoffmann, H.; and Lear, P. 2022. Tackling Misinformation with HAMLET (Human and Machine in the Loop Evaluation and Training). Technical report, Logically.
- BBC News. 2020. ‘Hundreds dead’ Because of Covid-19 Misinformation.
- Beatty, S. 2022. 6 Simple Search Tips: Lessons Learned from the Scopus Webinar.
- Botnevik, B.; Sakariassen, E.; and Setty, V. 2020. BRENDA: Browser Extension for Fake News Detection. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2117–2120. ACM.
- Bucchi, M. 2017. Credibility, Expertise and the Challenges of Science Communication 2.0. *Public Understanding of Science*, 26(8): 890–893.
- Chen, Z.; and Freire, J. 2020. Proactive Discovery of Fake News Domains from Real-Time Social Media Feeds. In *Companion Proceedings of the Web Conference 2020*, 584–592. ACM.
- Ciccone, K.; and Vickery, J. 2015. Summon, EBSCO Discovery Service, and Google Scholar: A Comparison of Search Performance Using User Queries. *Evidence Based Library and Information Practice*, 10(1): 34–49.
- Das, A.; Liu, H.; Kovatchev, V.; and Lease, M. 2023. The State of Human-Centered NLP Technology for Fact-Checking. *Information Processing & Management*, 60(2): 103219:1–103219:25.
- Freiling, I. 2019. Detecting misinformation in online social networks: A think-aloud study on user strategies. *SCM Studies in Communication and Media*, 8(4): 471–496.
- Google. 2022. Search Quality Rater Guidelines: An Overview.
- Graves, L. 2017. Anatomy of a Fact Check: Objective Practice and the Contested Epistemology of Fact Checking. *Communication, Culture & Critique*, 10(3): 518–537.
- Guo, B.; Ding, Y.; Yao, L.; Liang, Y.; and Yu, Z. 2020. The Future of False Information Detection on Social Media: New Perspectives and Trends. *ACM Computing Surveys*, 53(4): 68:1–68:36.
- Gupta, V.; Beckh, K.; Giesselbach, S.; Wegener, D.; and Wirtz, T. 2021. Supporting Verification of News Articles with Automated Search for Semantically Similar Articles. In *Proceedings of the 2021 Workshop Reducing Online Misinformation through Credible Information Retrieval*. CEUR Workshop Proceedings.
- Hasanain, M.; and Elsayed, T. 2022. Studying Effectiveness of Web Search for Fact Checking. *Journal of the Association for Information Science and Technology*, 73(5): 738–751.
- He, L.; and He, C. 2022. Help Me #DebunkThis: Unpacking Individual and Community’s Collaborative Work in Information Credibility Assessment. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2): 413:1–413:31.
- Ipsos MORI. 2022. Ipsos MORI Veracity Index.
- Jabiyev, B.; Pehlivanoglu, S.; Onarlioglu, K.; and Kirda, E. 2021. FADE: Detecting Fake News Articles on the Web. In *Proceedings of the 16th International Conference on Availability, Reliability and Security*, 15:1–15:10. ACM.
- Juneja, P.; and Mitra, T. 2022. Human and Technological Infrastructures of Fact-Checking. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2): 418:1–418:36.
- Kevin, V.; Högden, B.; Schwenger, A., Claudia and Şahan; Madan, N.; Aggarwal, P.; Bangaru, A.; Muradov, F.; and Aker, A. 2018. Information Nutrition Labels: A Plugin for Online News Evaluation. In *Proceedings of the 1st Workshop on Fact Extraction and VERification*, 28–33. ACL.
- Krieg, S. J.; Schnur, J. J.; Marshall, J. D.; Schoenbauer, M. M.; and Chawla, N. V. 2020. Pandemic Pulse: Unraveling and Modeling Social Signals During the COVID-19 Pandemic. *Digital Government: Research and Practice*, 2(2): 19:1–19:9.

- La Barbera, D.; Roitero, K.; and Mizzaro, S. 2022. A Hybrid Human-In-The-Loop Framework for Fact Checking. In *Proceedings of the 6th Workshop on Natural Language for Artificial Intelligence*, 4:1–4:10. CEUR Workshop Proceedings.
- Lewandowski, D. 2015. Evaluating the Retrieval Effectiveness of Web Search Engines Using a Representative Query Sample. *Journal of the Association for Information Science and Technology*, 66(9): 1763–1775.
- McClure Haughey, M.; Povolito, M.; and Starbird, K. 2022. Bridging Contextual and Methodological Gaps on the “Misinformation Beat”: Insights from Journalist-Researcher Collaborations at Speed. In *Proceedings of the 2022 ACM SIGCHI Conference on Human Factors in Computing Systems*, 244:1–244:15. ACM.
- Micallef, N.; Armacost, V.; Memon, N.; and Patil, S. 2022. True or False: Studying the Work Practices of Professional Fact-Checkers. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW1): 127:1–127:44.
- Nakov, P.; Corney, D.; Hasanain, M.; Alam, F.; Elsayed, T.; Barrón-Cedeño, A.; Papotti, P.; Shaar, S.; and Da San Martino, G. 2021. Automated Fact-Checking for Assisting Human Fact-Checkers. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence*, 4551–4558. IJCAI.
- Nguyen, A. T.; Kharosekar, A.; Krishnan, S.; Krishnan, S.; Tate, E.; Wallace, B. C.; and 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*, 189–199. ACM.
- Palmer, A. 2020. *Scientific Facts in the Space of Public Reason: Moderate Idealization, Public Justification, and Vaccine Policy Under Conditions of Widespread Misinformation and Conspiracism*. Ph.D. thesis, Bowling Green State University.
- Papagiannopoulou, E.; and Tsoumakas, G. 2020. A Review of Keyphrase Extraction. *WIREs Data Mining and Knowledge Discovery*, 10(2): e1339:1–e1339:45.
- Pidikiti, S.; Zhang, J. S.; Han, R.; Lehman, T.; Lv, Q.; and Mishra, S. 2020. Understanding How Readers Determine the Legitimacy of Online News Articles in the Era of Fake News. In *Proceedings of the 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 768–775. IEEE.
- Rose, M. E.; and Kitchin, J. R. 2019. pybliometrics: Scriptable Bibliometrics Using a Python Interface to Scopus. *SoftwareX*, 10.
- Sun, Y.; Qiu, H.; Zheng, Y.; Wang, Z.; and Zhang, C. 2020. SIFRank: A New Baseline for Unsupervised Keyphrase Extraction Based on Pre-Trained Language Model. *IEEE Access*, 8: 10896–10906.
- Tandoc Jr., E. C.; Ling, R.; Westlund, O.; Duffy, A.; Goh, D.; and Wei, L. Z. 2018. Audiences’ Acts of Authentication in the Age of Fake News: A Conceptual Framework. *New Media & Society*, 20(8): 2745–2763.
- Teyssou, D. 2019. *Applying Design Thinking Methodology: The InVID Verification Plugin*, 263–279. Springer.
- Vo, N.; and Lee, K. 2018. The Rise of Guardians: Fact-Checking URL Recommendation to Combat Fake News. In *Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, 275–284. ACM.
- Wang, J.; and Yu, B. 2021. News2PubMed: A Browser Extension for Linking Health News to Medical Literature. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2605–2609. ACM.
- Weld, G.; Glenski, M.; and Althoff, T. 2021. Political Bias and Factualness in News Sharing Across More Than 100,000 Online Communities. *Proceedings of the International AAAI Conference on Web and Social Media*, 15(1): 796–807.
- Zhou, X.; and Zafarani, R. 2020. A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. *ACM Computing Surveys*, 53(5): 109:1–109:40.