

ISSN: 2150-8402

Volume 14, Number 3, August 2022



Journal of Intelligent Learning Systems and Applications



ISSN: 2150-8402



<https://www.scirp.org/journal/jilsa>

Journal Editorial Board

ISSN Print: 2150-8402 ISSN Online: 2150-8410

<https://www.scirp.org/journal/jilsa>

Editor-in-Chief

Prof. Wei-Chiang Hong

Asia Eastern University of Science and Technology, Chinese Taipei

Editorial Advisory Board

Prof. Rob Law

Hong Kong Polytechnic University, China

Dr. Steve S. H. Ling

University of Technology Sydney, Australia

Prof. Yi Pan

Georgia State University, USA

Prof. Xin Xu

National University of Defense Technology, China

Editorial Board

Prof. Ala Al-Fuqaha

Western Michigan University, USA

Prof. George Angelos Anastassiou

Department of Mathematical Sciences, The University of Memphis, USA

Prof. Tao Ban

National Institute of Information and Communications Technology, Japan

Prof. Mokhtar Beldjehem

University of Ottawa, Canada

Dr. Rifai Chai

University of Technology, Sydney, Australia

Dr. Kit Yan Chan

Curtin University, Australia

Dr. Longbiao Chen

Xiamen University, China

Dr. Ying Guo

Commonwealth Scientific and Industrial Research Organization, Australia

Prof. Kao-Shing Hwang

National Sun Yat-sen University, Chinese Taipei

Prof. Fakhreddine Karray

University of Waterloo, Canada

Dr. Michail G. Lagoudakis

Technical University of Crete, Greece

Dr. Hak-Keung Lam

King's College London, UK

Prof. Shutao Li

Hunan University, China

Prof. Yiguang Liu

Sichuan University, China

Prof. Kin Huat Low

Nanyang Technological University, Singapore

Dr. Vincenzo Pacelli

University of Foggia, Italy

Prof. George Panayiotakis

University of Patras, Greece

Dr. Phyo Phyo San

Nanyang Technological University, Singapore

Prof. Steven Weidong Su

University of Technology Sydney, Australia

Prof. Theodore B. Trafalis

The University of Oklahoma, USA

Dr. Jue Wang

Chinese Academy of Sciences, China

Prof. Xuesong Wang

China University of Mining and Technology, China

Prof. Simon X. Yang

University of Guelph, Canada

Prof. Shi Ying

Wuhan University, China

Prof. G. Peter Zhang

Georgia State University, USA

Dr. Yudong Zhang

Columbia University, USA

Dr. Yi Zhao

Harbin Institute of Technology Shenzhen Graduate School, China

Prof. Wei Zhong

University of South Carolina Upstate, USA

Prof. Dingxuan Zhou

City University of Hong Kong, China

Prof. Antanas Zilinskas

Institute of Mathematics and Informatics, Lithuania

Table of Contents

Volume 14 Number 3

August 2022

Towards Immunizing Infodemic: Comprehensive Study on Assessing the Role of Artificial Intelligence and COVID-19 Pandemic

M. Roshanaei, G. Sywulak.....25

Journal of Intelligent Learning Systems and Applications (JILSA)

Journal Information

SUBSCRIPTIONS

The *Journal of Intelligent Learning Systems and Applications* (Online at Scientific Research Publishing, <https://www.scirp.org/>) is published quarterly by Scientific Research Publishing, Inc., USA.

Subscription rates:

Print: \$89 per copy.

To subscribe, please contact Journals Subscriptions Department, E-mail: sub@scirp.org

SERVICES

Advertisements

Advertisement Sales Department, E-mail: service@scirp.org

Reprints (minimum quantity 100 copies)

Reprints Co-ordinator, Scientific Research Publishing, Inc., USA.

E-mail: sub@scirp.org

COPYRIGHT

Copyright and reuse rights for the front matter of the journal:

Copyright © 2022 by Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>

Copyright for individual papers of the journal:

Copyright © 2022 by author(s) and Scientific Research Publishing Inc.

Reuse rights for individual papers:

Note: At SCIRP authors can choose between CC BY and CC BY-NC. Please consult each paper for its reuse rights.

Disclaimer of liability

Statements and opinions expressed in the articles and communications are those of the individual contributors and not the statements and opinion of Scientific Research Publishing, Inc. We assume no responsibility or liability for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained herein. We expressly disclaim any implied warranties of merchantability or fitness for a particular purpose. If expert assistance is required, the services of a competent professional person should be sought.

PRODUCTION INFORMATION

For manuscripts that have been accepted for publication, please contact:

E-mail: jilsa@scirp.org

Towards Immunizing Infodemic: Comprehensive Study on Assessing the Role of Artificial Intelligence and COVID-19 Pandemic

Maryam Roshanaei, Gregory Sywulak

The Pennsylvania State University Abington College, Abington, USA
Email: mur45@psu.edu

How to cite this paper: Roshanaei, M. and Sywulak, G. (2022) Towards Immunizing Infodemic: Comprehensive Study on Assessing the Role of Artificial Intelligence and COVID-19 Pandemic. *Journal of Intelligent Learning Systems and Applications*, 14, 25-41.

<https://doi.org/10.4236/jilsa.2022.143003>

Received: July 1, 2022

Accepted: August 28, 2022

Published: August 31, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Artificial Intelligence (AI) technologies have intentionally and unintentionally been used to spread false information on all different types of subjects. Throughout the COVID-19 pandemic, there was a pool of different information that was being presented to the public, a lot of it contradicting one another. False information spreads regardless of whether there is intent to mislead or misinform whereas AI is not able to decipher what type of information it is pushing to the public is correct and what is not. This mass spread of information through online platforms has been coined as an Infodemic where it is considered a massive volume of information, both online and offline. It includes deliberate attempts to disseminate false information to undermine the public health response and advance alternative agendas of groups or individuals. An infodemic can be incredibly dangerous to society greatly affecting the ability of communities, societies, and countries to control and stop the pandemic due to the abundance of different information in combating the health crisis. This article assesses and evaluates the role of Artificial Intelligence (AI) technologies in helping to spread disinformation during the COVID-19 pandemic. It reviews and evaluates the information curation in modern media, the relationship between AI and disinformation, and the challenges of disinformation campaigns. It further outlines the impact of social media platforms on infodemic and their influence in spreading disinformation during the COVID-19 pandemic. This article analyzes several data mining studies that used different machine learning techniques to identify the influence of disinformation tactics on the COVID-19 pandemic associated with the Twitter platform. It further continues exploring the investigation of the number of influential tweets, the type of users, the levels of credibility of URLs, and the type and effect of social media bots. Finally, the authors assess and conclude how disinformation is widely prevalent throughout social media dur-

ing the COVID-19 pandemic as well as illustrate the surveys that categorize the prevalence of users involved in the conversation about disinformation separated by country including the percentage of users posting tweets and retweeting news URLs, and the future work in combating the rapid disinformation campaigns and their ethical implication impact.

Keywords

Artificial Intelligence, Infodemic, Disinformation, COVID-19 Pandemic

1. Introduction

The fourth industrial revolution¹ has shaped the integration and interaction of different information and communication technologies in merging cyber, physical, and social infrastructure. Recent progress in the development of Artificial Intelligence (AI) has been immense and is growing exponentially to tackle technical and socio-economic challenges. AI solutions have enormous potential to accelerate the progress and influence the evolution of smart, sustainable, and guarantee stability to develop the communities' standards. AI is a technology that is here to make our lives easier to perform tasks more efficiently than any human could do however, on the other side of the token it is not advanced enough yet to be completing tasks that involve reasoning, real-world knowledge, and social interaction [1]. Therefore, AI technologies can potentially contribute intentionally and unintentionally in spreading false information on all different types of subjects. Throughout the COVID-19 pandemic, there was a vast information that was being presented to the public, a lot of it contradicting one another. False information spreads regardless of whether there is intent to mislead or misinform whereas AI is not able to decipher what type of information it is pushing to the public is correct and what is not. It is broadly defined the information activities² into three distinct types: Mis-, Dis-, Malinformation (MDM). Misinformation refers to unintentionally created and shared to misleads without any intent of causing harm. Whereas, disinformation is deliberately created, crafted, and shared with the intent to deceive, mislead, harm, or manipulate a recipient, social group, organization, or country. The malinformation is derived from fact to weaponize sabotage, harm, manipulate, or mislead. MDM manipulates and shapes public opinion and undermines trust in the authenticity of the information. Consequently, these activities are used to cause chaos, confusion, and division in democracy as well as national cohesiveness. Since the beginning of the COVID-19 pandemic, social media was the primary platform that spread information on how to be safe and connected. However, it turned into a global problem as AI algorithms lack deciphering and differentiating between factual and false information. This ushered in what has now been coined by the WHO (World Health

¹<https://www.foreignaffairs.com/articles/2015-12-12/fourth-industrial-revolution>.

²CISA—<https://www.cisa.gov/mdm>.

Organization)³ as an infodemic (information pandemic). Infodemic is defined as “an overabundance of information—some accurate and some not—that makes it hard for people to find trustworthy sources and reliable guidance when they need it”⁴. Although infodemic is amplified and crafted by social media with the help of AI to spread Mis-, Dis-, Misinformation, only disinformation with a manipulative character has been considered a threat that must be tackled. In the initial stages of COVID-19 pandemic, a survey [2] was conducted among residents of Central Pennsylvania about their knowledge of the pandemic and how to properly slow the spread of the virus. It concluded that the residents were worried and overwhelmed by the abundance of information and the mixed messages. The point of this survey was to determine if the public had adequate knowledge about the coronavirus, where they were receiving their information, how willing they were to comply with public health recommendations, and their level of trust in different information sources. This survey found that 4 out of 5 people had adequate knowledge about the pandemic but many of those people who had that knowledge still commented that they feel they do not know enough about how to manage symptoms of COVID-19. About 43% of participants said that government websites were their most trusted source of information, and 27% said that news media was their most trusted source. It was also noted that some people had negative feelings about how the pandemic was portrayed in the news followed by others being concerned that people had politicized the response to this pandemic to push their agenda. This in turn helped create this distrust of the information provided by the executive branch of the government for some groups. One of the three key worries that participants had and were able to write about in their short response area was that they had a feeling of information overload and did not know what to believe. The massive spread of COVID-19 has been accompanied by an infodemic since there is a demand for information on this disease. This demand for information about COVID-19 was the perfect opportunity for disinformation [3] to get mixed in with accurate information. Some of this disinformation is easy to spot because of how ludicrous it sounds but others can be extremely detrimental since the average person could not discern if what they were reading was true or not. This need for information has developed a spike in news consumption around the world. In a survey performed by *GlobalWebIndex*⁵, it was found that because of the pandemic 67% of people that were surveyed have watched more news coverage. This allowed people to be more easily fooled by information that sounds legitimate but is misleading. The main objective of this article is to focus on and contribute a comprehensive study in assessing the role of AI technologies in helping to spread disinformation during the COVID-19 pandemic as well as reviewing the information curation in modern media, the relationship between AI and disinformation, and the challenges of disinformation campaigns. The con-

³WHO—<https://www.who.int/>.

⁴Infodemic—<https://iris.paho.org/handle/10665.2/52052>.

⁵<https://www.gwi.com/coronavirus>.

tribution outlines the impact of social media platforms on infodemic and their influence in spreading disinformation during the COVID-19 pandemic. It employs different studies and analyses the influence of disinformation during the COVID-19 pandemic by evaluating recent data mining studies comprising tweets associated with COVID-19, a comparison of several types of Twitter accounts in proponents versus opponents of COVID-19 vaccinations, negative tweets about COVID-19 vaccinations, the source of news, and age distribution using the popular source to get information. It continues with an investigation of the number of influential tweets, the type of users, the levels of credibility of URLs, and the type and effect of social media bots. Finally, it evaluates a study that goes further and categorizes the prevalence of users involved in the conversation about disinformation separated by country and the percentage of users posting tweets and retweeting news URLs.

2. Information Curation in Modern Media

One interesting thing from the media is that a video game talked about some of the issues we are facing today back in 2001 when it was released [4]. This game dives into issues about curators of information, AI, as well as false information. First, when talking about curators of information this game states that throughout history there have always been people who write and record history as well as people who only pass down certain information about history. Now today, those curators cannot be any group of people because there is simply too much information out there to be able to go through it all and definitively say what is true and what is not. That is where AI comes in to enable sifting through all the information that is available to use on the internet to be able to do what no human can do. This is exactly what this video game does inside its storyline. They find out that there is a need for someone or something to organize and verify the information that is out there, so they created this AI to do just that. Now, in the video game, they take a more science fiction route where AI then becomes sentient so, that is where the story breaks off and the discourse ends for this topic. But it is very interesting that even back in 2001 the game creator was talking about an issue that we would face 20 years later when the mass amounts of information about COVID-19 would be available to the public without any person being able to curate and verify what is true and what is not.

3. Relationship between AI and Disinformation

In the early stages of AI, the first iterations were going for a type of human intelligence. This is very similar to how it is portrayed in popular media such as movies where it has conscious thought and has free will. Their idea [5] when doing this was to be able to reverse engineer the thought process of humans and be able to teach that to machines and have it automated. That is vastly different from how AI is created today. Today, AI is created to in theory surpass itself. Consequently, they are no longer trying to replicate human thought but teaching

it the rules they need to follow and then telling it to learn how to complete the task as efficiently as possible. It is in turn learning from the mistakes it makes and looking for the best possible solution to solve the problem they were assigned.

4. Challenges of Disinformation Campaigns

Campaigns of disinformation are not new when it comes to controlling public opinion whether it be based on elections or other divisive public topics. There are recorded instances back in 1796 where the French government⁶ engaged in heavy disinformation and fear tactics to tilt the scale towards the election of Thomas Jefferson who was the pro-French candidate. Back in 2016, there were similar tactics employed but on a completely different scale compared to what was achieved back in 1796. During the 2016 election AI technology was used to promote foreign interests as well as disinformation. This is called computational propaganda that uses big data and machine learning about the users to allow it to manipulate public opinion on a bigger scale as well as more efficiently. It used bots to target individuals or different demographics that are known to be more susceptible to the messaging that they are pushing. The foreign power [6] used AI bots and social media to push its agenda. This shows that information can be used as a weapon when trying to influence public opinion as well as the outcomes of some events like elections. There are similar things happening today from 2020 through the present day where there are groups, domestic, or foreign powers using social media to push disinformation about COVID-19 as a whole or just strictly about the vaccination. The key problem is to address how to combat the war on information and find a faster way of sifting through all of the information that is posted on their platform. At the same time, it is needed to ensure what to censor because the social media platforms such as Twitter and Facebook are open forums [7] so they do not take responsibility for what is posted on their platform. If they end up going too far and censoring discourse rather than obvious attempts to misinform or disinform the public then they will take away the basic right of freedom of speech on their platform. There is a very fine line [8] that they need to label and be able to clearly define what is disinformation and what might just be a discussion between two opposing ideologies. Content is catered for the specific person based on their likes and interests and this is the concept of digital personalization where AI fits into that picture. The AIs can then be used maliciously based on the data they receive from personal accounts. When they find out your preferences then they can tailor what message they will use on you to make sure it is the most effective. Social media spreads information like wildfire. This was especially evident during the COVID-19 pandemic. Whenever there is new information or new guidelines released it then

⁶DeConde, A. (1957) Washington's Farewell, the French Alliance, and the Election of 1796. *The Mississippi Valley Historical Review*, 43(4), 641-658.

<https://doi.org/10.2307/1902277>

spreads very quickly to millions of people within mere minutes. Although AI is what is helping this infodemic to spread so quickly it is also what is being used to slow it down. Facebook has set into place an AI that actively monitors and removes posts of disinformation. From March 1st through November 9th in 2020 [9] there were more than 180 million pieces of content that were checked with a third-party fact-checker that was flagged or completely removed from their platform. The COVID-19 crisis is not all just about the dangers of the disease that is spreading but, it is also about the spread of harmful disinformation throughout social media. Social Media bots [10] have been observed to be about twice as active as they have been for other crises or events that have happened such as elections. What these bots have done throughout these different events is spread all sorts of information about the hot topic that the world is watching very closely. This information is not always accurate and there is a plethora of bots designed to strictly promote disinformation. Although AI is at the forefront of this issue it could also be the solution to the problem. AI is an incredible technology that can control the spread of certain types of information based on how they were created. It can help disinformation spread like wildfire, but it can also be put into place to decipher the true information from the false on these large social media platforms. This advanced AI technology [10] can be given more parameters as to how they will plan their attack on spreading this disinformation such as who to attack when to attack, and what would be the best way for maximum effectiveness and exposure. Some ways that it can spot these malicious spreaders is first to go through social media and do a bot spot. This has an AI sift through numerous accounts on a specific social media platform and be able to tell what accounts are real people and what accounts are these bots that are spreading disinformation. During lockdown more people are spending time online as well as on social media. These domains [11] are open forums in which you can post topics such as conspiracy theories, false and misleading information. These types of information thrive on these types of forums since there is not much moderation. This allowed hateful extremists to spread their disinformation and engage a larger audience. These groups or individual people use bots to spread their message because it is quicker and more efficient than physically doing it yourself. Consequently, this use of bots has made it difficult to track and neutralize the source of the disinformation. Tech companies have tried to start using good bots to combat the massive spread of disinformation but they are still a long way off from having full control over these falsehoods that continue to spread.

5. Impact of Social Media Platforms on Infodemic

During the COVID-19 Pandemic, this was the first time that technology such as social media was used to spread information on how to be safe and connected. This turned into a problem because at the same time there is disinformation being spread that the public does not know how to decipher the difference between the fake information from the factual ones. Social media has a plethora of in-

formation that has been shared through it. Although this information is readily accessible it is not readily reliable. When you read or watch something on social media about COVID-19 you must take extra steps to confirm the information that has been presented in that post. The problem is that many people do not take that extra step when reading about COVID-19 and they just take the post for face value believing it to be true [12]. This post then gets shared and spread to other people in their social group extending the reach of this disinformation.

6. Influence of Disinformation on COVID-19 Data Mining Studies

In a data mining study [13], Twitter was queried looking for tweets associated with COVID-19. Tweets that were excluded were duplicates that came from retweeting or sharing as well as academic source tweets. This left 13,596 non-academic tweets about COVID-19. The rest of the tweets are nonspecific. The methodology used in this study was to extract all Tweets and hashtags related to COVID-19 using keywords including “COVID”, “COVID-19”, “corona”, “coronavirus”, “positive”, “test”, “tested”, “feel”, “I”, “we”, “my”, “us”, “our” using Twitter programming interface and performing data mining with R version 3.2.3 and subsequently Python version 3.4.2. In this study, it was found that tweets related to COVID-19 in nonacademic settings primarily contain unverifiable information or disinformation altogether. The study found the different topics 1) disinformation on the relationship between influenza infection and COVID-19 infection 2) belief of getting the influenza vaccine can lead to a positive COVID-19 test 3) belief of 5G networks and regional outbreaks and infections of COVID-19. **Figure 1** illustrates the tweets containing disinformation categorized by the topic.

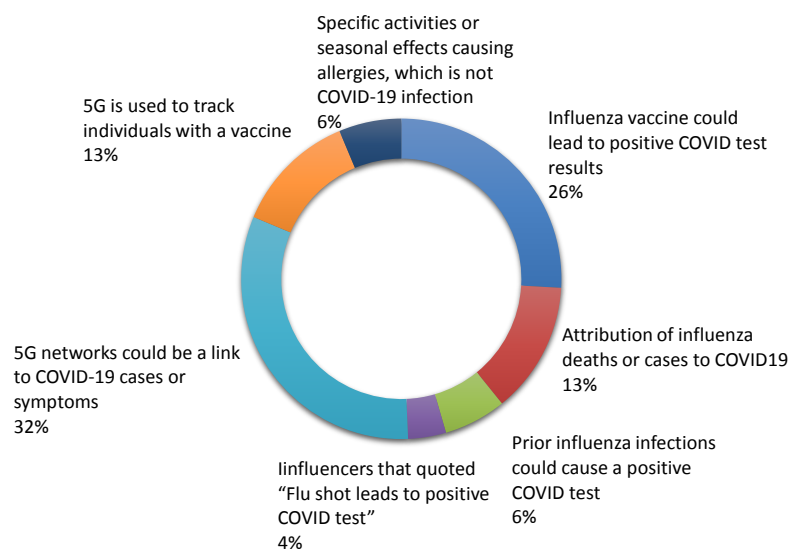


Figure 1. Tweets containing disinformation categorized by the topic.

Another study performed by researchers from [14] investigated the several types of accounts that were proponents of COVID-19 vaccines versus accounts that were opponents of COVID-19 vaccines. This study used a sample size of 2000 accounts to get their findings. The study is used Botometer [15] with majority of accounts had low scores with indication of low likelihood of automation. They broke down these accounts into three main groups: vaccine opponents, vaccine proponents, and others. Others did not have a clear stance on vaccinations. About a quarter of all accounts studied posted primarily about vaccines which means that they do not post about anything else besides vaccinations whether they are pro-vaccine or against it. **Figure 2** illustrates the percentage of accounts in proponents versus opponents of COVID-19 vaccinations.

The other study [16] investigated the major thematic areas of sentiments towards a COVID-19 vaccine. They collected a total of 1,286,659 tweets from the timeline of July 19, 2020, to August 19, 2020. This study employed topic modeling technique including an unsupervised machine learning technique to screen tweets for inclusion and identify the common topics; Latent Dirichlet Allocation (LDA) to determine the comparison between the relevant tweets with the irrelevant tweets containing COVID-19 vaccination keywords; and finally ran a sensitivity test with multiple values to determine the number of clusters [17] [18]. From that, 20 separate topics were defined and a total of 4868 tweets were organized in these categories based on the highest probability of belonging to this specific cluster. They were then labeled positive, negative, neutral, or irrelevant. Out of the 4868 tweets chosen 1306 of them were labeled negative and these are the categories of tweets assigned to negative. These tweets contained statements like that the vaccine is not necessary because the survival rate is high as well as the worries of the side effects of a vaccine being developed that quickly. **Figure 3** illustrates collected negative tweets about COVID-19 vaccinations.

In the United States, there are more than 8 out of 10 people get their news from a smartphone computer, or tablet sometimes or often. [19] employed an online surveyed U.S. adults used methodology⁷ with the representation of the U.S. adult population by gender, race, ethnicity, partisan, affliction, education and other categories. There are provided a set of questions⁸ for this analysis, along with responses and its methodology⁹. The chances that people have encountered disinformation that they did not even know was false are remarkably high. The total statistics are shown in **Figure 4**.

The different age groups vary wildly on what platforms online they receive their information from. In the youngest group from 18 - 29 years old, social media is the most popular platform for receiving information but as the groups get older in an age that number goes down and the most popular place to receive

⁷Pew Research center's American Trends Panel (ATP):

<https://www.pewresearch.org/our-methods/u-s-surveys/the-american-trends-panel/>.

⁸https://www.pewresearch.org/journalism/wp-content/uploads/sites/8/2021/01/PJ_2021.01.12_News-and-Social-Media_TOPLINE.pdf.

⁹<https://www.pewresearch.org/journalism/2021/01/12/news-consumption-on-social-media-methodology/>.

information is on news websites or apps. The Age Distribution using the most popular place online to get information is shown in **Figure 5**.

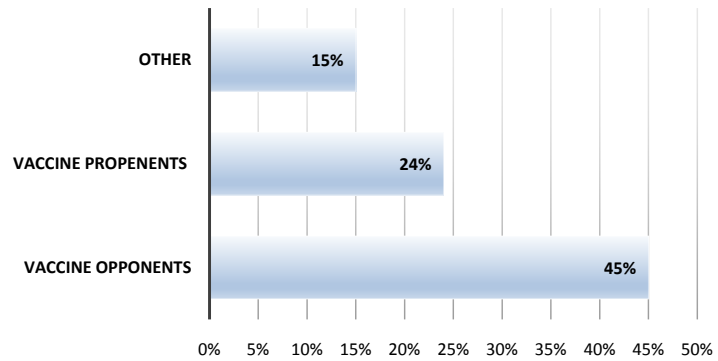


Figure 2. Percentage of proponents and opponents who accounts for COVID-19 vaccinations.

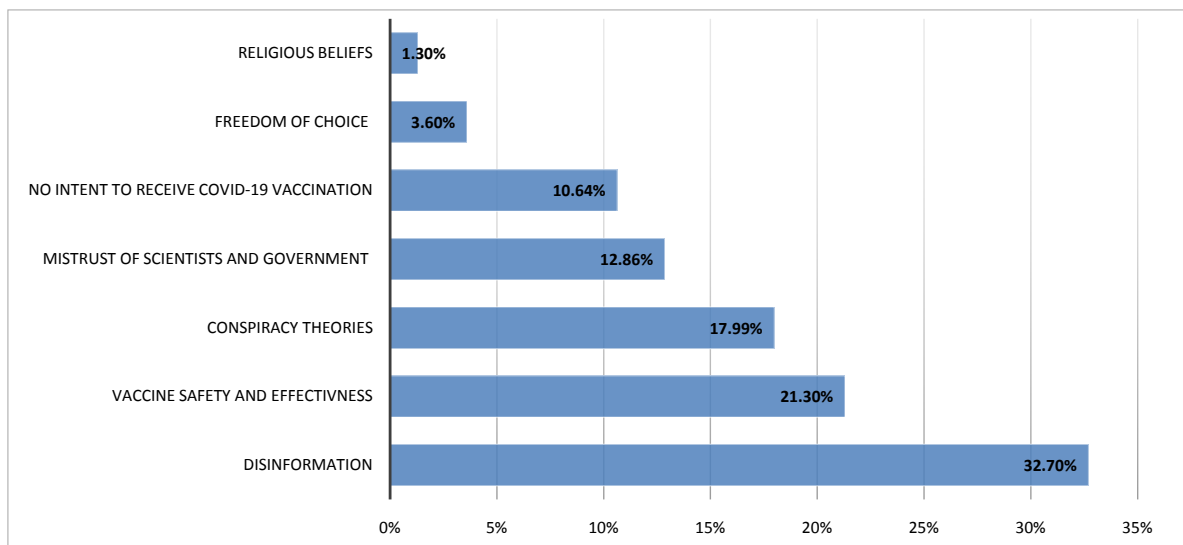


Figure 3. Negative tweets about COVID-19 Vaccination.

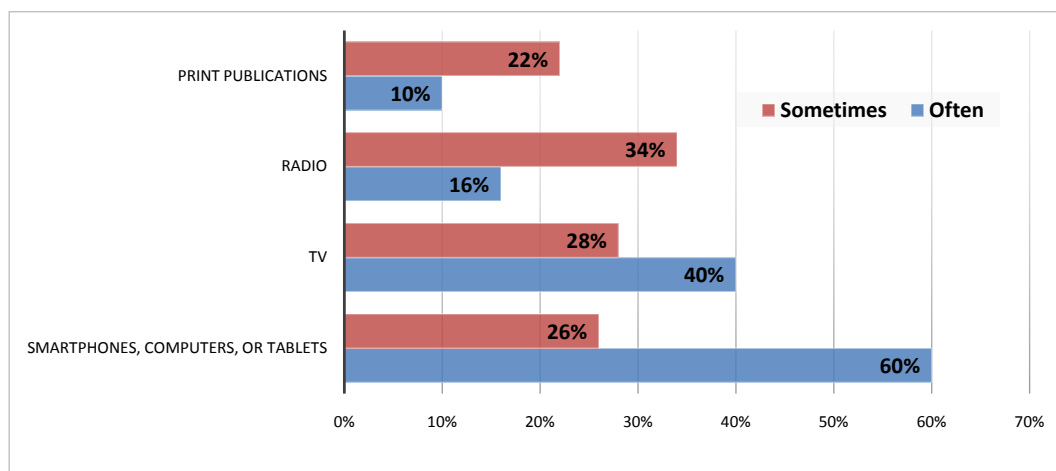


Figure 4. Source of news.

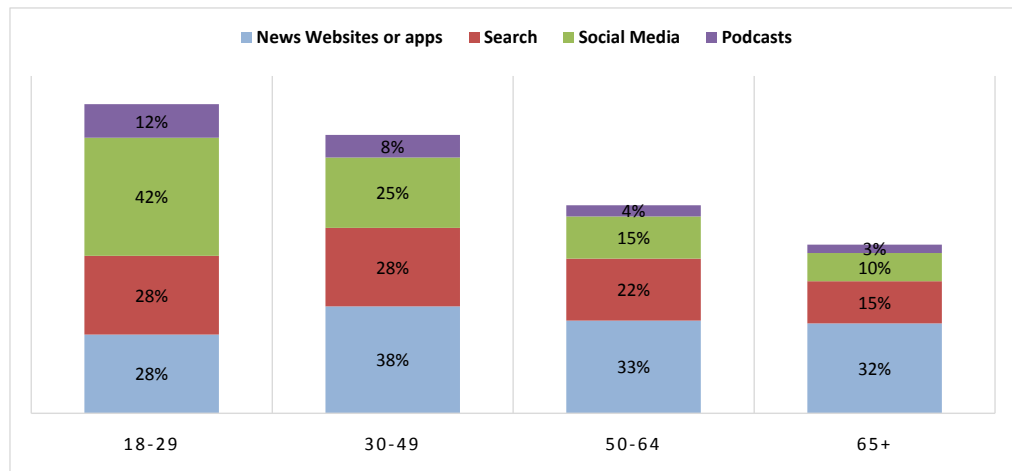


Figure 5. Age Distribution using the most popular place online to get information.

People are 70% more likely to retweet or share disinformation on different social media platforms so it will reach the first 1500 people six times faster than that of its counterpart [20]. Labeling these posts as disinformation comes with a downside because then after those posts without these warnings on them is all perceived to be true even though that might not be the case. One way that AI has played a role in the divide between groups of people in the United States is that it will show you more about topics that you have interacted with so if you interact with a post with disinformation, you will be more likely to see more posts that are along those same lines. There has been a trend in the rise of disinformation over the past few years. You can see whenever there is a crisis or a major event that has a lot of coverage worldwide there is a surge of disinformation on social media platforms as well as other online forums. From [21], there are data sets from the years 2015-2018 that were compiled as a list of sites producing disinformation by combining six lists [22]-[27]. In these data sets, they show that leading up to the 2016 election there was a rise in fake news websites as well as disinformation as a whole.

7. Influential Tweets Categories

There was another study completed in June of 2020 [28] that uses a sample size of over 12 million tweets and from there those tweets were then categorized by the type of user, and the content, as well as popularity. This study addressed the type of users who send influential tweets during the global health emergency event, identifying the type of users who discuss disinformation stories and global network with low credibility information. The used methodology approach to monitoring COVID-19 conversation on Twitter with selected keywords including “coronavirus”, “coronavirus”, “wuhan virus”, “wuhavirus”, “2019nCoV”, “NCoV”, “NCoV2019” and identify user’s; location, social identity and political orientation. First starting they categorized the type of users that spread the most influential tweets and the influence of social media bots as shown in **Table 1**.

Table 1. Influential tweets, the user categories, and social media bots.

	Percentage of types of users who spread influential tweets	Percentage of users who have the 10,000 most influential tweets	Percentage of bots
Regular Users	96.44%	30.73%	25.98%
News Agencies	1.12%	32.91%	12.38%
News Reporters	1.18%	18.04%	8.79%
Government officials	1.05%	19.94%	48.29%

The data showed that the spread of influential Tweets is just regular users at 96.44%. The next highest category is news reporters at 1.18%. Subsequently, when talking about the people who spread the most influential tweets it is shown that they are just regular people who do not have any sort of credibility about what they post. This can become a huge problem if those people decide to start using their platforms to sway the masses in whatever way they want them to. Contrary to that the next categorization of tweets is taking the top 10,000 most influential tweets out of the sample size and using that as the new sample. When only the top 10,000 tweets are considered, it changes the dataset. The most influential type of tweeter is regular people who still hold 25.98% of tweets inside the top 10,000 most influential. However, with that being said the other types of users have a higher percentage of tweets in the top 10,000 when compared to their presence overall. Even though news agencies were only 1.12% of the total data size of influential tweets they have 12.38% of the top 10,000 most influential. This is showing that even though they have a small amount in the total data sample they also have some of the most influential tweets out of all the categories. Now, when looking at these the question might be asked when AI and disinformation come into play with these most influential tweets. First, when looking at the prevalence of AI this study examined each user, and from the behavior of the account as well as the content that has been posted, they determine whether or not it can be labeled as a bot account. Subsequently, as shown in **Table 1**, the highest percentage of bots was found in news agencies followed closely by regular users. News agencies having bots run some of their accounts makes sense because as said above these accounts that are set up can be used to do repeatable tasks and sharing the news daily to Twitter is something that can be automated. However, for regular users there raises the question is why there is such a high percentage of bots inside that type. Moving on to the prevalence of disinformation in this data set they then investigated what types of users cite fake news sites or even discuss disinformation storylines. To explain this the first thing that needs to be understood is how they classify these websites. They used a color system dependent on what the content is on that specific fake news website. The three classifications they used were black site, red site, and orange site. They just called sites with credible information to be “real sites”. Black sites are

websites that exclusively post fabricated stories. Red sites are ones that spread falsehoods using a flawed editorial process. Orange sites are where it was harder to tell if they were using a flawed editorial process but at the same time, there were still falsehoods on their page. These different charts are showing the number of tweets split up by websites that each user group is referencing in their tweets. Regular users are by far the most prominent type of user on Twitter so they are going to be in the top percentile of whatever type of news site it might be, but the interesting thing happens when the percentage of bot accounts comes into play. **Figure 6** is outlining the percentage of bots found in the same tweets that were referencing these separate URLs. One thing that can be seen from **Figure 6** is that bots are mostly using black news websites when referencing websites for their information. As you go down the list the percentage of bots decreases when getting to the information that is getting more “true.” This shows that the bots’ jobs are meant to pull from certain websites to gain a following or to gain more attention since it was stated that disinformation spreads six times faster than its true counterpart.

This study goes further and categorizes the prevalence of these tweets filled with disinformation by country. The countries that they used were English-speaking countries since all these tweets in the data set are written using English (EN). As depicted in **Figure 7** the most prominent English-speaking countries and their percentage of users that posted either “fake news” or “real news” URLs in their tweets.

The most telling from **Figure 8** is that the United States is the only one where the percentage of “fake news” URLs posted is higher than the percentage of “real news” URLs posted. On the other hand, the percentage of users that retweeted URLs from either “fake news” or “real news” sources is sorted by country. Once again, the United States is the only English-speaking country that had a higher percentage of users retweeting “fake news” articles over “real news” articles. **Figure 8** illustrates the percentage of users involved in the conversation about disinformation separated by country.

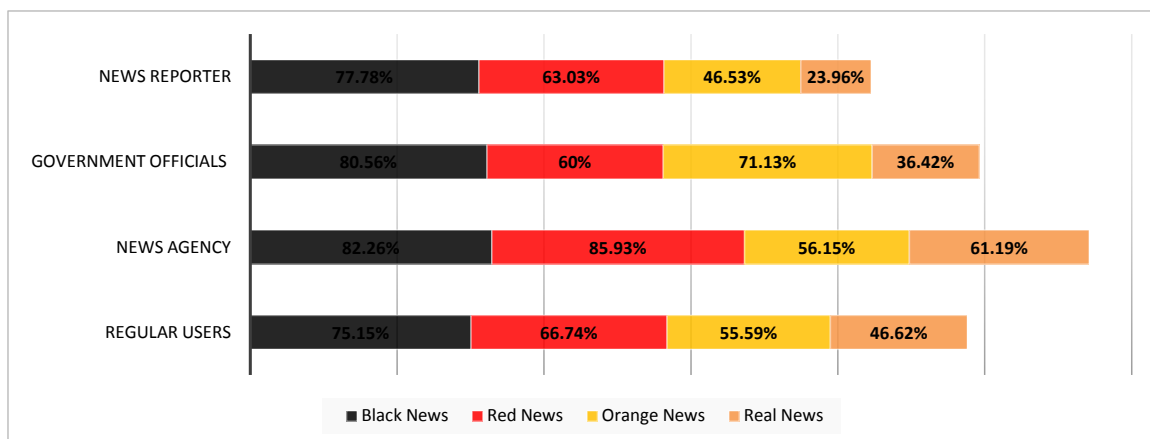


Figure 6. Number of tweets, type of users, and the levels of credibility of URLs (Black, Red and Orange).

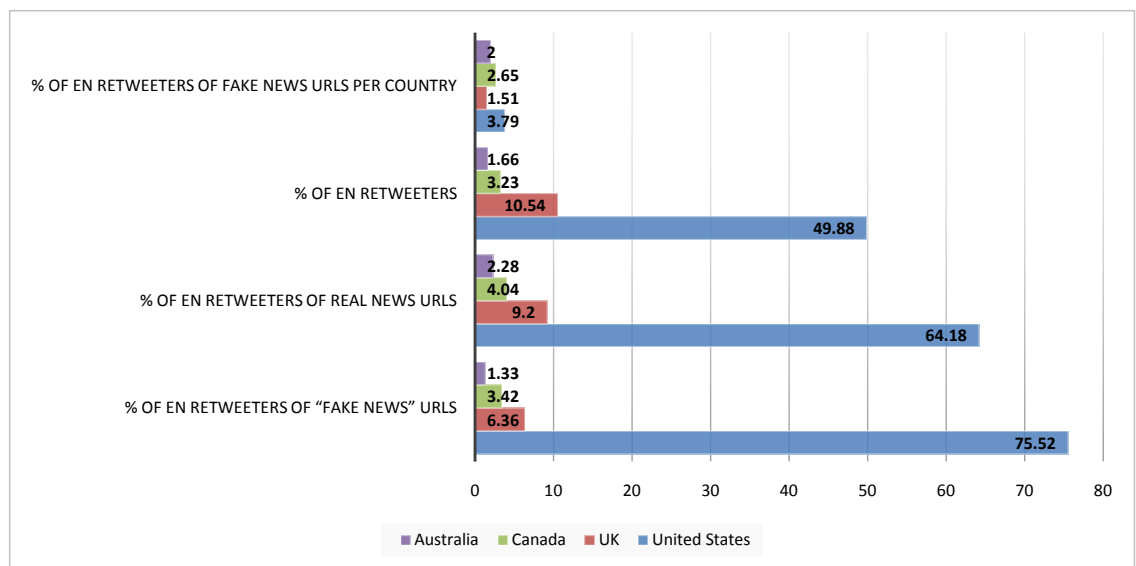
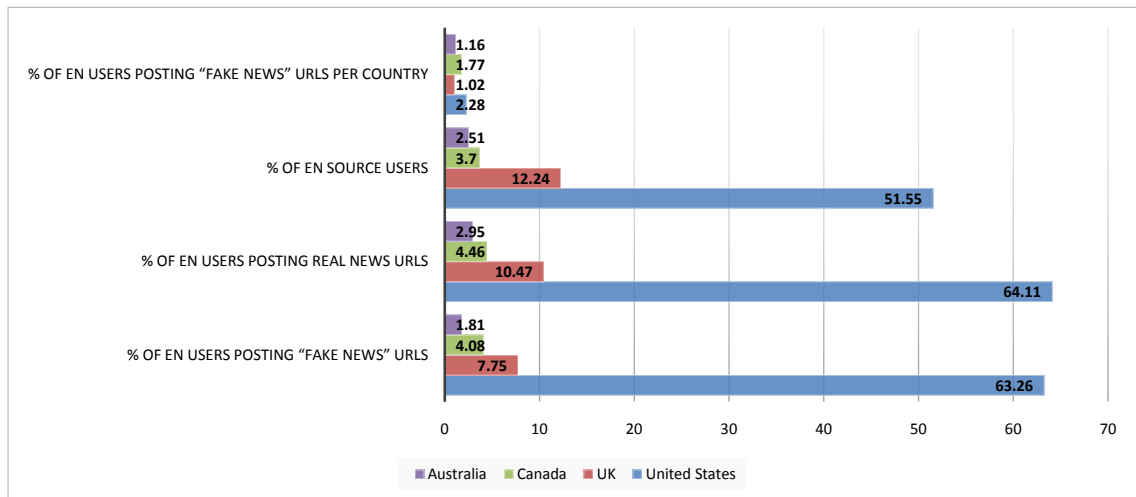


Figure 7. Percentage of users posting tweets and retweeting with news URLs.

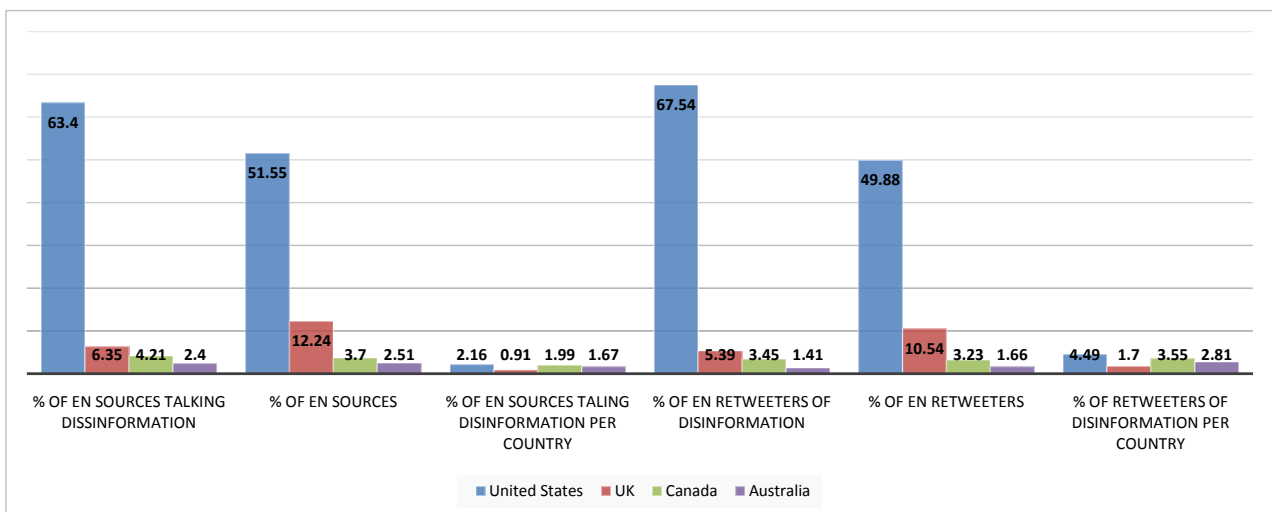


Figure 8. Users involved in the conversation of disinformation per country.

With the observation of all these data, it becomes clear that the United States first off has the most posts and conversations about COVID-19 whether it comes from “fake news” sources or “real news” sources. Being said they also have a higher percentage of posts, retweets, or talking about disinformation when you compare it to their percentage of posts, retweets, or talking about real news sources. This shows that tweets about disinformation in the United States are more prevalent than tweets about real news sources.

8. Conclusion

Disinformation is widely prevalent throughout social media in today’s society. The rise of COVID-19 has only given more fuel to the fire for the spread of disinformation. A key part of fighting the pandemic is fighting against the disinformation that has spread like wildfire. This article showed with the data analysis that the reason that disinformation has been able to spread so quickly is because of AI and how it pushes posts out to users on its perspective platform. The articles’ data analyses of the disinformation influence on COVID-19 evaluated the recent data mining studies comprise of tweets associated with COVID-19, a comparison of several types of Twitter accounts in proponents versus opponents of COVID-19 vaccinations, negative tweets about COVID-19 vaccinations, the source of news, and age distribution using the popular source to get information. It evaluated a study that goes further and categorized the prevalence of users involved in the conversation about disinformation separated by country and the percentage of users posting tweets and retweeting news URLs. These studies showed that the groups of people continually spread disinformation about the COVID-19 vaccination which has caused dissent among those groups and they remain unvaccinated. This combination of AI and users sharing these posts filled with disinformation has allowed the Infodemic to progress this far. By looking at data sets from Twitter it was able to be determined that disinformation is at an all-time high on social media and there are no signs of slowing down. It was clear that the vast number of accounts that are spreading disinformation on Twitter are bot accounts that spew out set information as well as interact with other bot accounts that have similar content. There is no set way of how to deal with this Infodemic however Twitter as well as other social media companies have tried taking steps to combat the spread of disinformation. With that being said they still need to come up with a better way to combat this mass amount of information because if this problem keeps ensuring it will only drive the divide between societies to become larger than it already is. In parallel to this article’s assessment and evaluation, the authors recommend that ethical implication is another key point for future work in combating the rapid disinformation campaigns. It requires further studies that should carefully look into public consensus where still several unresolved questions raise regarding the First Amendment protection and human dignity versus determining the universal description of how to judge the desirable or undesirable in society, how to distinguish the generated contents as

harmful with regards to the importance of the free speech protection.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Pretz, K. (2021) Stop Calling Everything AI, Machine-Learning Pioneer Says Michael I. Jordan Explains Why Today's Artificial-Intelligence Systems Aren't Intelligent. IEEE Spectrum.
<https://spectrum.ieee.org/stop-calling-everything-ai-machinelearning-pioneer-says>
- [2] Van Scoy, L.J., Miller, E.L., Snyder, B., Wasserman, E., Chinchilli, V.M., Zgierska, A.E., Rabago, D., Lennon, C.L., Lipnick, D., Toyobo, O., Ruffin 4th, M.T. and Lennon, R.P. (2021) Knowledge, Perceptions, and Preferred Information Sources Related to COVID-19 Among Central Pennsylvania Adults Early in the Pandemic: A Mixed Methods Cross-Sectional Survey. *The Annals of Family Medicine*, **19**, 293-301.
<https://doi.org/10.1370/afm.2674>
- [3] Fleming, N. (2020) Coronavirus False Information, and How Scientists Can Help to Fight It. *Nature*, **583**, 155-156. <https://doi.org/10.1038/d41586-020-01834-3>
<https://www.nature.com/articles/d41586-020-01834-3>.
- [4] Derrat, M. (2019) The Most Profound Moment in Gaming History.
https://www.youtube.com/watch?v=jYBod0ge3Y&ab_channel=MaxDerrat
- [5] Dick, S. (2019) Artificial Intelligence. *Harvard Data Science Review*, No. 1.1.
<https://doi.org/10.1162/99608f92.92fe150c>
- [6] Kreps, S. (2020) The Role of Technology in Online False Information.
<https://www.brookings.edu/wp-content/uploads/2020/06/The-role-of-technology-in-online-misinformation.pdf>
- [7] Twitter Safety (2021) Updates to Our Work on the COVID-19 Vaccine False Information.
https://blog.twitter.com/en_us/topics/company/2021/updates-to-our-work-on-COVID-19-vaccine-misinformation
- [8] Lian, P. (2019) AI Is Helping Spread False Information Faster. How Can We Deal with That? *ITU News*.
<https://aiforgood.itu.int/ai-is-helping-spread-misinformation-faster-how-can-we-deal-with-that/>
- [9] Meta AI (2020) Here's How We're Using AI to Help Detect False Information.
<https://ai.facebook.com/blog/heres-how-were-using-ai-to-help-detect-misinformation/>
- [10] Slapakova, L. (2021) Towards an AI-Based Counter-Disinformation Framework.
<https://www.rand.org/blog/2021/03/towards-an-ai-based-counter-disinformation-framework.html>
- [11] Cox, K., Ogden, T., Jordan, V. and Paille, P. (2021) COVID-19, Disinformation and Hateful Extremism Literature Review Report. Prepared for the Commission for Countering Extremism (CCE). RAND Europe, Cambridge.
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/993841/RAND_Europe_Final_Report_Hateful_Extremism_During_COVID-19_Final_accessible.pdf
- [12] Goel, A. and Gupta, L. (2020) Social Media in the Times of COVID-19. *JCR: Journal*

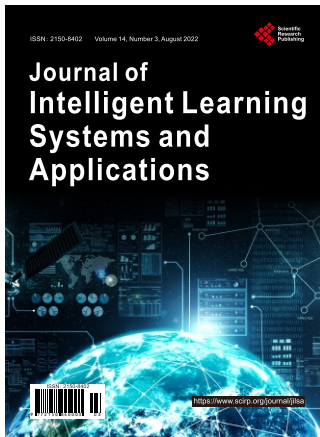
of *Clinical Rheumatology*, **26**, 220-223.

<https://doi.org/10.1097/RHU.0000000000001508>

- [13] Krittanawong, C., Narasimhan, B., Virk, H.U.H., Narasimhan, H., Hahn, J., Wang, Z., *et al.* (2020) Misinformation Dissemination in Twitter in the COVID-19 Era. *The American Journal of Medicine*, **133**, 1367-1369.
<https://doi.org/10.1016/j.amjmed.2020.07.012>
- [14] Jamison, A.M., Broniatowski, D.A., Dredze, M., Sangraula, A., Smith, M.C. and Quinn, S.C. (2020) Not Just Conspiracy Theories: Vaccine Opponents and Proponents Add to the COVID-19 'Infodemic' on Twitter. *Harvard Kennedy School Misinformation Review*. <https://doi.org/10.37016/mr-2020-38>
- [15] Davis, C.A., Varol, O., Ferrara, E., Flammini, A. and Menczer, F. (2016) BotOrNot: A System to Evaluate Social Bots. *Proceedings of the 25th International Conference Companion on World Wide Web (WWW'16 Companion)*, Montréal, 11-15 April 2016, 273-274. <https://doi.org/10.1145/2872518.2889302>
- [16] Nuzhath, T., Tasnim, S., Sanjwal, R.K., Trisha, N. F., Rahman, M., Mahmud, S., *et al.* (2020, December 11). COVID-19 Vaccination Hesitancy, Misinformation and Conspiracy Theories on Social Media: A Content Analysis of Twitter Data. *SocArXiv*. <https://doi.org/10.31235/osf.io/vc9jb>
- [17] Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003) Latent dirichlet allocation. *Journal of machine Learning Research*, **18**, 993-1022.
- [18] Surian, D., Nguyen, D.Q., Kennedy, G., Johnson, M., Coiera, E. and Dunn, A.G. (2016) Characterizing Twitter Discussions about HPV Vaccines Using Topic Modeling and Community Detection. *Journal of Medical Internet Research*, **18**, e232.
<https://doi.org/10.2196/jmir.6045>
- [19] Shearer, E. (2021) More than Eight-in-Ten Americans Get News from Digital Devices. Pew Research Center. <https://pewrsr.ch/2MZqns7>
- [20] Brown, S. (2020). MIT Sloan Research about Social Media, False Information, and Elections. MIT Sloan.
<https://mitsloan.mit.edu/ideas-made-to-matter/mit-sloan-research-about-social-media-misinformation-and-elections>
- [21] Allcott, H., Gentzkow, M. and Yu, C. (2019) Trends in the Diffusion of False Information on Social Media. *Research & Politics*, **6**, Article ID: 205316801984855.
<https://doi.org/10.1177/2053168019848554>
- [22] Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B. and Lazer, D. (2019) Fake News on Twitter during the 2016 U.S. Presidential Election. *Science*, **363**, 374-378. <https://doi.org/10.1126/science.aau2706>
- [23] PolitiFact (2017) Politifact's Guide to Fake News Websites and What They Peddle. PolitiFact.
<http://www.politifact.com/punditfact/article/2017/apr/20/politifacts-guide-fake-news-websites-and-what-they/>
- [24] Silverman, C. (2016) Here Are 50 of the Biggest Fake News Hits on Facebook from 2016. BuzzFeed News.
<http://www.buzzfeednews.com/article/craigsilverman/top-fake-news-of-2016>
- [25] Silverman, C., Singer-Vine, J. and Vo, L.T. (2017) In Spite of the Crackdown, Fake News Publishers Are Still Earning Money from Major ad Networks. BuzzFeed News.
<http://www.buzzfeednews.com/article/craigsilverman/fake-news-real-ads>
- [26] Guess, A., Nyhan, B. and Reifler, J. (2018) Selective Exposure to Misinformation: Evidence from the Consumption of Fake News during the 2016 U.S. Presidential Campaign. Working Paper, European Research Council, Brussels.

<https://about.fb.com/wp-content/uploads/2018/01/fake-news-2016.pdf>

- [27] Schaedel, S. (2017) Websites That Post Fake and Satirical Stories. FactCheck. <http://www.factcheck.org/2017/07/websites-post-fake-satirical-stories/>
- [28] Huang, B. and Carley, K.M. (2006) Disinformation and Misinformation on Twitter during the Novel Coronavirus Outbreak. ArXiv: abs/2006.04278.



Journal of Intelligent Learning Systems and Applications

ISSN Print: 2150-8402 ISSN Online: 2150-8410

<https://www.scirp.org/journal/jilsa>

The *Journal of Intelligent Learning Systems and Applications (JILSA)* is a peer reviewed international journal with a key objective to provide the academic and industrial community a medium for presenting original cutting-edge research related to intelligent learning systems and their applications. JILSA invites authors to submit their original and unpublished work that communicates current research on intelligent learning systems both in the theoretical and methodological aspects, as well as various applications in real-world applications.

Papers are invited on the topics including, but not limited to:

- Approximate Dynamic Programming
- Autonomic Computing
- Autonomous Learning Systems
- Bio-Inspired Learning Method
- Clustering
- Cyber-Physical Systems
- Evolutionary Computation
- Filter Bank
- Fuzzy and Rough Set
- General Theory on Intelligent Learning Systems
- Intelligent Applications
- Learning Control Systems
- Multi-Agent Learning
- Multiple Models Strategy
- Neural Networks
- Pattern Recognition Based on Learning Techniques
- Predictive Control
- Recommendation System
- Reinforcement Learning
- Robotics
- Smart Grid Techniques and Applications
- Social Network Analysis
- Statistical Learning Theory
- Supervised Learning
- Time-Frequency Representation
- Unsupervised Learning
- User Behavior Modelling
- Wavelet Theory

Editor-in-Chief

Prof. Wei-Chiang Hong

Asia Eastern University of Science and Technology, Chinese Taipei

Notes for Intending Authors

Submitted papers should not have been previously published nor be currently under consideration for publication elsewhere. Paper submission will be handled electronically through the website. All papers are refereed through a peer review process. For more details about the submissions, please access the website.

Website and E-Mail

<https://www.scirp.org/journal/jilsa>

E-Mail: jilsa@scirp.org

What is SCIRP?

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

What is Open Access?

All original research papers published by SCIRP are made freely and permanently accessible online immediately upon publication. To be able to provide open access journals, SCIRP defrays operation costs from authors and subscription charges only for its printed version. Open access publishing allows an immediate, worldwide, barrier-free, open access to the full text of research papers, which is in the best interests of the scientific community.

- High visibility for maximum global exposure with open access publishing model
- Rigorous peer review of research papers
- Prompt faster publication with less cost
- Guaranteed targeted, multidisciplinary audience



**Scientific
Research
Publishing**

Website: <https://www.scirp.org>

Subscription: sub@scirp.org

Advertisement: service@scirp.org